Putting the *Folk* Back in Folk Psychology:
The Social, Cultural, and Moral Character of Folk Psychology

Craig Lambert Roxborough

A Dissertation Submitted to
the Faculty of Graduate Studies
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

Graduate Program in Philosophy
York University
Toronto, Ontario

April 2016

Abstract

This project provides a critical analysis of how philosophers have traditionally characterized and understood folk psychology, what I call the traditional construal of folk psychology. On the basis of a growing and diverse body of empirical evidence regarding how folk psychology is deployed *in situ* by the folk, I argue that we must reject this traditional construal and embrace a new understanding of folk psychology.

Two crucial and related assumptions regarding the function of folk psychology have been taken for granted in the contemporary debate regarding the mechanisms that underlie our folk psychological competence. More specifically, philosophers have assumed that the primary function of folk psychology is to explain and predict behaviour and that folk psychology is a quasi-scientific enterprise, whereby these goals are shared with science.

However, an empirical investigation into how the folk use folk psychology reveals a different story. Instead, we find that we often fail to accurately explain and predict behaviour and that our folk psychological discourse functions so as to satisfy a number of social, cultural, moral goals or purposes. In this way, there are a number of normative functions to our folk psychological practices.

These observations, I argue, are inconsistent with the traditional construal of folk psychology and therefore warrant that we take a skeptical position with respect to the traditional construal and re-examine the assumptions that underlie it. However, I argue that attempts to reconcile the traditional construal with the empirical evidence are unlikely to be successful and that the right way forward is to reject the tradition and embrace a new understanding of folk psychology. The traditional construal, I argue, is too narrow in scope to adequately address this empirical evidence and we must embrace a more comprehensive and empirically informed understanding of folk psychology that can embrace the normative interests that shape and drive this practice. While much work will remain, this new landscape will be a significant departure from the traditional construal and we will leave this project having put the *folk* back into folk psychology.

For my family

Acknowledgements

I must begin by thanking my mother, Elizabeth Roxborough, for her love, support, and encouragement throughout my education as I regularly changed directions and discovered my ultimate path and passion. But most importantly, I must thank her for teaching me (in Grade 3!) how to take notes, turn those notes into a first draft, and then polish and revise my work into a final draft. At the time I had no idea just how invaluable this lesson would be. My committee, Kristin Andrews, Henry Jackman, and Muhammad Ali Khalidi, this project would not have been possible without your support and critical eye. Your work inspired me and you regularly pushed me to evaluate my ideas and rethink my work in ways that challenged me to be better than I thought I could be. In particular, Kristin, I am indebted to you for your endless patience as I took months and years to produce and revise chapters and your insight into my work helped me see my project in new and better ways. Thank you to all my friends, co-workers, and managers who encouraged me along the way and wouldn't let me quit. In particular, thank you to Jeremy Petch for his seemingly endless support and the endless shame he would have brought upon me if I gave up and to Julia Brassolotto for the encouragement, insight into my project, and random text messages of support while I was writing that helped keep me going. Thank you as well to my son Charlie, whose conception accelerated the completion of this project and thank you to Joanne and Ed Gallagher for helping to take care of my family (and my lawn!) while I worked. Lastly, I am unbelievably grateful to my wife Jen. Your undying support, patience, and encouragement allowed me to finish this work. All the missed evenings, weekends, and summer vacations; I hope this project makes you proud and I cannot thank you enough.

Table of Contents

Chapter One: Introduction

1.1 – Introduction

To say that folk psychology has played a significant role in contemporary philosophy of mind would be an understatement. From metaphysics of mind to artificial intelligence, from consciousness studies to developmental psychology, folk psychology has played an essential role. It can act as a limiting condition on the possibility of the reduction of mental states to physical states, it can offer guiding principles in designing artificial cognitive systems, and in practice it has shaped entire domains of research in psychology and has significantly influenced the way we approach investigations in animal cognition research. The importance of folk psychology to our understanding of mind cannot be understated and philosophers have spilled a great deal of ink in a seemingly endless debate regarding the possible mechanisms that underlie our folk psychological competence.

In this project I provide a critical analysis of contemporary philosophy of mind as it pertains to the characterization and understanding of folk psychology. Specifically, there is an assumption[1] that underlies much of the philosophical debate around folk psychology that I will argue fails to accord with a growing body of empirical evidence regarding how and why the folk actually engage in folk psychological discourse. In effect, the current and dominant interpretation or understanding of folk psychology, what I will call the traditional construal, is not an accurate portrayal of folk psychology as used by the folk and fails to capture important elements of this real world practice. Importantly, this project shifts the focus away from armchair theorizing about folk psychology towards an empirically motivated and informed analysis that puts the *folk* back into folk psychology.

[1]Really a set of assumptions that I group into one category.

This project is predominantly negative or skeptical in nature. That is, I am primarily concerned with motivating a rejection of the current theoretical landscape regarding folk psychology in order to motivate the need for change moving forward. In effect, I wish to reject the status quo and encourage investigators interested in folk psychology to re-examine some firmly held assumptions regarding what folk psychology is designed to do and how the folk deploy it in practice. This is a hefty task. It requires developing an argument against, at minimum, thirty years of research and theorizing about folk psychology that is also firmly rooted in key historical works in the philosophy of action.

Providing a clear articulation of this negative project and encouraging philosophers to reject some firmly held assumptions is a valuable task that deserves full treatment and defence prior to engaging in the next steps and exploring the new landscape of understanding. In effect, we cannot move forward with this discussion until we accept that we've been going down the wrong path for quite a long time and for now, this change in perspective warrants significant attention and argumentation. This approach is not new, nor unfamiliar. In order to develop a new understanding of some issue or phenomenon we must often break down significant epistemological barriers.[2] More specifically, those barriers that are erected by the view we are trying to dispose of. But unfortunately, these barriers often prevent or at least make it more difficult for us to move forward. This has been noted by many, including the French philosopher and historian Gaston Bachelard[3] and then, perhaps most famously, by the philosopher Thomas S. Kuhn (1962). Our current thinking may often pose challenges to progress and ultimately, sometimes knowledge production simply means rejecting the past. Of course, this process may

[2] Thank you to Julia Brassolotto for inspiring this particular view of my project. See also Brassolotto, Raphael & Baldeo (2014) for a further discussion of epistemological barriers, and in particular, how they serve as a barrier for addressing the social determinants of health among public health professionals in Ontario.
[3] See Tiles (1984) for translation and discussion.

not be linear, it is often disjointed, and sometimes it requires significant breaks from the past that may involve a wholesale rejection of the previous theory. While this description of knowledge production is often accepted as true in science, it applies equally well in philosophy and this project is an example of this process. I will advance an argument in favour of fracturing our current thinking so that we can break from the current discourse and move ourselves forward, re-orienting our focus in the direction of truth as it pertains to our understanding of folk psychology as a real-world, observable, and purposeful practice.

While the bulk of this project will be focused on advancing this critical argument and leaving much of the new theoretical work to future endeavours, we will explore and learn much about the new landscape that flows from this project. In particular, as we explore empirical evidence that ultimately supports the rejection of the traditional construal of folk psychology, much will be revealed about the new landscape and how we should understand the folk practice of folk psychology. Moreover, in demonstrating the limitations of current thinking regarding folk psychology, I will explore in detail some new approaches that are significant breaks from the tradition but that accord with the empirical observations made throughout this project. While not a definitive defense of any of these options, exploring these new conceptions of folk psychology will help to reveal much about how we can and should conceive of folk psychology moving forward.

In order to move the discussion forward, I will re-examine the status quo regarding folk psychology using an empirical lens. Once viewed through this lens we will come to see that the traditional assumptions regarding folk psychology are not consistent with how the folk use folk psychology in practice. This inconsistency, I will argue, creates challenges for those who accept these traditional assumptions. I will argue that the best way to reconcile this inconsistency is to

reject our current understanding as too narrow and fundamentally mistaken about the purpose of folk psychology, and to begin to build a new understanding of how and why the folk use folk psychology in practice. Much work will remain following this project, but demonstrating that long-held assumptions regarding folk psychology are mistaken is a significant fracturing of our thinking that will require significant re-orientation in our thinking as we construct a new understanding of folk psychology. This is an important shift in thinking and given the interaction between our current understanding of folk psychology and other issues in philosophy, it is likely that this shift in thinking will not be benign. However, there is a long road ahead before we get to this point and a map to guide us will be helpful.

1.2 – A Roadmap

To begin, in Chapter Two I will examine the current state of the philosophy of mind literature as it pertains to folk psychology in order to reveal assumptions that underlie how folk psychology is traditionally construed. Through a brief exploration of the history of folk psychology and the traditional debates regarding folk psychology, I will reveal two crucial and related assumptions philosophers have made about folk psychology. More specifically, I will demonstrate that those interested in folk psychology have traditionally conceived of folk psychology as an explanatory and predictive enterprise. Put another way, that the primary function of folk psychology is to facilitate the explanation and prediction of behaviour. Additionally, I will demonstrate that this construal of the function of folk psychology also portrays folk psychology as being quasi-scientific in nature where folk psychology and science share the same goal of explaining and predicting behaviour and that these explanations and predictions are thought to be causal in nature. This set of assumptions, I will argue, is not consistent with empirical evidence regarding

folk psychology and must ultimately be rejected in favour of a more comprehensive and empirically informed account of folk psychology.

However, before critically evaluating the traditional construal of folk psychology I will devote Chapter Three to defending the empirical lens adopted by and utilized throughout this project. In particular, I will defend the use of an empirical lens specifically as it relates to an evaluation of the traditional construal of folk psychology. While I suspect that the application of an empirical lens to this philosophical issue will be fairly non-controversial, because much of what will follow flows from the new experimental philosophy or "X-Phi" movement which has drawn significant criticism, a defense of this approach as it applies to folk psychology is warranted. Ultimately, I will argue that because the genuine investigator of folk psychology is engaged in a descriptive project with the goal of understanding a real-world practice of the folk, we must be prepared to utilize the empirical lens to evaluate our theoretical commitments. More specifically, we must ensure that our theoretical commitments are properly aligned with the practice itself and the best method for doing so is to empirically investigate the real-world practice itself.

In Chapters Four and Five I'll apply this empirical lens to folk psychology in order to reveal that folk psychology as deployed *in situ* by the folk is a significantly different enterprise than philosophers have traditionally thought it to be. While explanation and prediction have dominated our theorizing about why philosophers think the folk use folk psychology, the empirical evidence I will explore reveals that folk psychology is designed and used for entirely different purposes. In fact, what we find is that folk psychology often fails to facilitate the production of explanations and predictions of behaviour and instead, is often significantly sensitive to social, cultural and moral considerations. That is, when we engage in folk

psychological discourse we are often attempting to satisfy a number of social, cultural, and moral goals or purposes. In this way, the empirical evidence reviewed and collected throughout this project demonstrates that folk psychology as used by the folk has a normative dimension to it.

In Chapter Six, I'll first demonstrate that the assumptions of the traditional construal of folk psychology are inconsistent with the empirical evidence just reviewed and collected. More specifically, that the empirical evidence challenges the claim that folk psychology functions to facilitate the quasi-scientific explanation and prediction of behaviour. This warrants, I will argue, that we take a skeptical position with respect to the traditional construal of folk psychology and be prepared to re-examine the assumptions that underlie it and ultimately reconcile our theoretical commitments with the empirical evidence if we aim to maintain this position. In effect, the burden of proof has been shifted and the traditionalist must defend their assumptions. However, I will then argue that our best approach to reconciling this inconsistency is to embrace a new understanding of folk psychology that requires us to significantly re-think how we view and understand this practice. That is, the traditional construal and underlying assumptions are not well positioned to account for this empirical evidence and our best approach for moving forward is to reject the traditional construal as too narrow in scope and embrace a more comprehensive understanding of folk psychology that more accurately reflects how folk psychology is used in practice by the folk.

While much of this project is somewhat radical in nature, elements of it are not entirely new. Much of the empirical evidence that will be explored in the pages that follow is quite well known (see for example, Knobe, 2006b; Nisbett, 2003; Nisbett and Wilson, 1977). Despite this, the consequences of this evidence have not yet been fully explored or addressed by those interested in folk psychology. Additionally, other philosophers have already begun to question

the claim that folk psychology should be understood as a quasi-scientific explanatory and predictive enterprise (see for example, Andrews, 2012; McGeer, 2007; Zawidzki, 2008). Unfortunately, to my knowledge this empirical evidence and critical investigation of folk psychology have not been sufficiently integrated. This project aims to do precisely this and some of those philosophers who are critical of the traditional construal will be explored in the closing pages of this project as they offer viable options for how to best account for this empirical evidence in the new folk psychological landscape. In this way, while the project may at some points have a radical feel to it, this project aims to significantly change our understanding of folk psychology by providing a novel synthesis of sometimes familiar but disparate observations or arguments and support this endeavour with a novel analysis that drives this integration forward into a new landscape and new way of conceiving of folk psychology.

The roadmap then is straightforward. We begin with a familiar trek into contemporary philosophy of mind examining the traditional understanding of folk psychology and subsequent debates. This exploration is not simply to situate the project within the larger debate, but to also reveal the assumptions that underlie our understanding of folk psychology and their role in the debate as we know it. With the assumptions articulated, I will shift gears significantly, leaving the armchair behind and embracing the empirical realm. We start with a short defence of the importance of empirical research in philosophy, the philosophy of mind and in particular, any investigation of folk psychology and then we jump head first into an ocean of data and discover how the folk actually utilize folk psychology, revealing not just what we think happens, but what actually does happen in practice. I will then argue that this empirical evidence ultimately demands that we reject the assumptions regarding folk psychology and explore a new landscape where significantly different accounts of folk psychology will be offered and adjudicated. As the

project closes I will begin to clean up some of the dust that I have kicked up and begin to sketch out what this new landscape will look like, what it means for our understanding of folk psychology, and explore some of the implications this new landscape will have on other philosophical questions. In effect, I will leave this project having significantly changed the folk psychological landscape and having put the *folk* back into folk psychology.

Chapter Two: A Brief History of Folk Psychology

2.1 – Introduction

The negative project that unfolds in the following chapters is premised on the claim that the goal of folk psychology, traditionally construed, is to explain and predict behaviour and the related assumption that explanation and prediction are to be understood in quasi-scientific terms. While I think this claim about the prominent and traditional view we are about to explore is fairly obvious and generally widely accepted, because the project that follows is so critical of this assumption it is important to ensure that I'm not unfairly or inaccurately characterizing the view that I aim to criticize.

With this in mind, in this chapter I will defend the claim that folk psychology is traditionally viewed as an explanatory and predictive enterprise. In other words, that the primary function of folk psychology is to facilitate the explanation and prediction of behaviour of both ourselves and others. In addition, I will show that construing folk psychology as an explanatory and predictive practice positions folk psychology as being quasi-scientific[4] in nature. That is, first that folk psychology and science share the same goal of explaining and predicting behaviour, and second, that folk psychological explanations and predictions are generally thought to be causal in nature.

I will accomplish this goal through a targeted review of some of the key philosophers in the contemporary debate regarding folk psychology. The claim here is not that all philosophers

[4] I use the term quasi-scientific to acknowledge that folk psychological explanations and predictions may not have the formal structure typically associated with scientific explanations and predictions. That is, while it is true that a theoretical characterization of folk psychology might portray the nature and structure of folk psychology as being entirely analogous to scientific theories (e.g. Churchland, 1981; Lewis, 1972), this need not be the case. The primary analogy that is being drawn between folk psychology and science is simply that the goal of these projects is the same. Simply put, both science and folk psychology have the goal of explaining and predicting behaviour. Additionally and importantly, there is good reason (as we will see) to think that folk psychological explanations and predictions have been assumed to be causally accurate, capturing the true causes of behaviour. In this way, the alignment between folk psychology and science is even stronger.

conceive of folk psychology in this way,[5] but rather just to say that among those who engage in the traditional debate regarding folk psychology that has been ongoing for at least the last three decades, the assumptions just outlined are held in some form. Put another way, while there is first order disagreement about the precise composition of folk psychology and mechanisms that underlie this competence, there is agreement at a higher level of analysis and it is this agreement about the core assumptions that I will call the traditional construal of folk psychology.

Through this targeted review we will be forced to explore some of the precise details of this traditional debate. In particular, the debate over the precise cognitive mechanisms that allow us to make folk psychological explanations and predictions. While I will wade into some of the details of this discussion I will be silent on the merits of the opposing views as this stage of the project is only indirectly related to this broader debate regarding how it is that we engage in folk psychological practices.

Rather, the focus of this project is one step removed from this debate, and in particular, one step prior to this debate as it focuses on the assumptions that underlie it. I am not interested in providing an account of the mechanisms that underlie our ability to engage in folk psychological discourse or attribute folk psychological concepts to ourselves and others. I am not interested in the merits of the opposing views or the merits of a hybrid approach to understanding this cognitive skill. Instead, in this project I am interested in and will focus on the growing side-debate that questions the assumptions that have been made about the purpose or function of folk psychology. As such, I am situated on the side-lines of this contemporary debate looking in from

[5] To suggest that all philosophers think of anything in the same way would be a bold and indefensible claim. In particular, with respect to folk psychology there has been a group of philosophers, which is now growing, that question the assumptions I have identified (see for example, Andrews, 2012; McGeer, 2007; Zawidzki, 2008). The view that I'm characterizing just happens to be a fairly widely held view that has dominated the literature for some time.

a position of skepticism and with a critical eye firmly focused on the assumptions that underlie it. This is important to disclose since my treatment of the debate below is quite superficial; I only wade into the details of the debate insofar as it helps to demonstrate the commitment these philosophers hold with respect to the assumptions I have identified above. However, the targeted review that follows and the admission I just made achieves the second goal of this chapter, to situate my project within the broader philosophical landscape and to ensure that my readers are not displeased as I skirt the challenging issues and questions that those I cite have tried to tackle since these are not the focus of this project.

2.2 – What is Folk Psychology?

What is folk psychology? This question is so broad, so general, and any answer we provide to this question is so likely to be influenced by our theoretical commitments that it is almost impossible to know how to begin to fairly provide an answer. Almost.[6]

At the most general level folk psychology is used by the lay person[7] (the folk) in order to navigate the social domain. As inherently social creatures we have the relatively[8] unique challenge of navigating not just the physical or material domain, the world of objects, but also the social domain, the world of subjects. As such we have evolved to navigate our way around not just material objects in our space, but also social subjects that we can interact with in entirely

[6] In order to start any investigation into what folk psychology is, we must venture at least some sort of guess to this answer so that we know what the object of our investigation is. My plan here is to start with an answer that is so general that almost any existing account of folk psychology would fall within the boundaries of my definition. Additionally, I have tried to be general to the point of avoiding any specific theoretical commitments in my reply to this question. While I'm sure some will object that this answer expresses some implicit theoretical commitments about folk psychology, I hope they will agree that these commitments are quite general and amenable to various interpretations.

[7] And philosophers when they're not being philosophers; psychologists when they're not being psychologists; neuroscientists when they're not being neuroscientists, etc.

[8] Of course other animals in social communities may be faced with similar challenges. Whether some other animals use something that resembles folk psychology is well beyond the scope of this project and a subject about which there is vigorous debate. The claim that *we* navigate a social domain is much clearer for us than it is for other animals.

different ways. Of course, this social domain is quite unlike the physical world and so we have evolved to utilize a different set of tools, concepts, and/or practices to work our way through this world, coordinating our own actions with those around us, figuring out what other people are going to do in response to our own actions, working together as a cooperative team to solve joint tasks or generally just working in close proximity to other thinking, moving, emoting beings.[9] Folk psychology is what allows us to do just this.

Regardless of the precise make-up of this practice, framework, network of tools and concepts, regardless of the precise mechanisms that allow us to deploy it in practice, one thing is clear: we navigate this domain with relative ease. Just as our understanding of physical objects allows us to navigate a room full of objects with ease (barring some clumsy moments), our understanding of the social domain, our folk psychology, allows us to navigate through a room of thinking moving subjects and if necessary interact with them in complex ways (barring some socially awkward interactions). In essence, we all seem to function well in our day-to-day interactions and the suggestion here is that our folk psychology is what allows us to succeed.

2.3 – Folk Psychology as Explanation and Prediction

So how exactly is it that folk psychology allows us to navigate the social domain? Put another way, how does folk psychology allow us to coordinate behaviour or walk the busy streets of downtown Toronto?[10] Traditionally speaking, philosophers have claimed that our folk psychological practice is simply the development of explanations and predictions of our own

[9] There are a multitude of possible purposes for folk psychology, these are just some of them.

[10] Some might suggest that this is more a matter of folk physics than folk psychology. There is some truth to this as in addition to being social beings, we are also physical things that move through space and need to be avoided. However, historically folk psychology has been invoked very broadly such that even navigating busy streets or driving a car could involve thinking about the mental states of others in order to achieve our goal of moving through a space full of thinking subjects who are making decisions as they move. That is, not only are we physical things moving through space, but we are thinking things moving through space and thinking things move differently than physical things and so folk psychology would be required to facilitate this navigation.

behaviour and the behaviour of others and that these explanations and predictions allow us to navigate the social domain. So, for example, our ability to avoid each other on the busy downtown streets of Toronto or my ability to coordinate with my wife while we cook dinner is the result of repeated explanations and predictions that we provide to ourselves, about ourselves and about others. This suggestion is fairly robust. What better way, we might think to coordinate behaviour or move around thinking subjects than to predict what they are going to do? What better way to come to understand the behaviour of others and ourselves than to explain it? Folk psychology, then, is the way in which we achieve this purpose.

While the central claim of this chapter is that the purpose of folk psychology is to provide explanations and predictions of behaviour, traditionally speaking this analysis of folk psychology has not come in isolation. That is, granting that folk psychology is used to explain and predict behaviour is only the first step, the second step requires an articulation of precisely how we achieve this goal. Traditionally, philosophers have construed folk psychological explanations and predictions as being comprised of mental states that we attribute to ourselves and others and combine in particular ways in order to explain and predict behaviour. In particular, philosophers have focused their attention on a specific observation regarding our[11] ability to attribute mental states such as beliefs and desires to one another, and they have claimed that our attribution of these unobservable mental states to ourselves and others allows us to explain and predict behaviour. That is, we go beyond just mere behaviour and consider the internal "goings on" of the person as a way to better understand why they are engaging in whatever behaviour they are doing, or to better predict what they are going to do.

[11] Typically adults, but also children as they mature in their psychological development. The debate regarding when children acquire the ability to attribute mental states to others is ongoing. See Onishi & Baillargeon (2005) for an argument that children as young as 15-months-old can predict behaviour on the basis of a true or false belief.

The precise nature[12] of the mental states that play an essential role in folk psychological explanations and predictions is, as should be expected, an ongoing debate. However, we can speak about them in general terms. Again, traditionally speaking philosophers interested in folk psychology have zeroed in on the class of mental states known as propositional attitudes.[13] In other words, the mental states that are most traditionally associated with folk psychology are those that are attitudinal in nature with respect to some specific proposition. The canonical propositional attitudes are beliefs and desires,[14] but of course one can hold other attitudes with respect to some proposition, but these two have been the focus of much of the discourse.[15] So one can have the attitude of belief with respect to some proposition, say, "it's raining outside". Analogously, one can have a desire regarding some proposition, say, "my having a beer". In either case there is a relationship between an attitude and some proposition. Tracing back to the observation that we attribute mental states to ourselves and others, examples of the attribution of

[12] For example, whether these mental states correspond to real physical states in the brain.
[13] While the class of mental states could in principle be quite large, philosophers have tended to focus their attention only on the class of propositional attitudes.
[14] As I will discuss in later chapters, there is good reason to think that folk psychology is actually composed of concepts well beyond beliefs and desires, and even the class of propositional attitudes.
[15] The history of thought in the philosophy of action may have something to do with this. If we go back to Aristotle's analysis of practical wisdom we see an early portrait of behaviour as resulting from the combination of various propositions about what one ought to do and the specifics of the case at hand (see for example, *De Anima 434a15-20; Nicomachean Ethics1147a25-30*). This purely rational view of human behaviour could be criticized on the grounds that reason alone is not enough to command action. Hence, we need some sort of desire or appetite to act in accordance with reason in order to produce action. Here we see the thinking of Hume filling in this gap. Hume argued that reason is inert and that in order for reason to produce action that some sort of appetite, effectively a desire, to act must be combined with reason (see *A Treatise of Human Nature*, especially 3.1.1). This idea was also picked up, in part, by Davidson (1963) who argued that every action is the product of a primary reason which is effectively the combination of reason and a desire for action. Throughout this train of thought, we can see a move from reason (which is effectively just a formal way of describing beliefs) to reason and desire working together. This focus on reason and this focus on desire or appetite, I think, has significantly influenced the way early philosophers interested in folk psychology conceived of and described our folk psychological practices. Given this history and focus, it is not too surprising to see philosophers interested in folk psychology zeroing in on just the propositional attitudes of beliefs and desires since (a) this accords with the history of thought in the philosophy of action, (b) there were good historical reasons for thinking that a belief and a desire together could cause an action, and (c) beliefs and desires are regularly attributed to people in our daily interactions and these are, quite frankly, two of the more interesting propositional attitudes or mental states that one could hold.

propositional attitudes could include “She believes it is raining outside” or “I want to drink a beer”.[16]

The suggestion that folk psychology involves the attribution of mental states such as beliefs and desires in order to explain and predict the behaviour of ourselves and others is not an unusual one and a moment of reflection on your own folk psychological practices should make this suggestion sound pretty familiar. Reflecting on my own experiences I can confidently say that I am always attributing beliefs and desires to those around me. Be it whether I’m trying to figure out if my wife actually does not want a gift for her birthday, whether my boss is impressed with me, whether the not quite old (but perhaps old enough) patron on the subway wants my seat, or what politicians and celebrities are thinking when they make strange decisions.

If we reflect on this practice even for a moment it really does appear, in the words of Jerry Fodor (1987), to be ubiquitous (p.3) and one that we all engage in all the time. Daniel Dennett (1987) even plausibly claims that the attributions of beliefs and desires to ourselves and others is unavoidable or nearly impossible to suppress (p.8), as something we simply cannot help but do when we observe someone engaged in some action. Propositional attitudes, then, really do seem to be components of our folk psychological discourse. I introduce propositional attitudes only because the traditional debate focuses on them and the discussion that follows will reference them regularly. I am not taking a particular position or stance on their nature or the role of they play in the attributive processes of folk psychology.

Returning now to the central claim of this chapter, that there is a prominent and traditional construal of folk psychology that is widely held and views the purpose of folk

[16] Of course the most formal, and perhaps proper instantiation of this propositional attitude would be “I want that I have a beer”. This exhibits the true attitude-proposition relationship that is indicative of the canonical propositional attitudes. I’ve chosen instead to use slightly less formal language which is closer to how we would in practice express this propositional attitude.

psychology to be the explanation and prediction of behaviour. This view is in fact widely held by those philosophers engaged in the early discourse over the nature of propositional attitudes and the cognitive mechanisms that underlie our ability to attribute them. This could not be made more obvious than by looking at the titles of two seminal papers by Paul Churchland: "Eliminative Materialism and the Propositional Attitudes" (1981) and "Folk Psychology and the Explanation of Human Behavior" (1989). The former is a critique of folk psychology through a naturalistic critique of propositional attitudes, while the latter continues this assault on folk psychology by looking specifically at an alternative explanatory framework for understanding the causes and nature of human behaviour. Combined they clearly indicate a commitment to the idea that propositional attitudes are central to folk psychology and that folk psychology is an explanatory enterprise. Moreover, Churchland clearly endorses the view that folk psychology is an explanatory and predictive enterprise when he states that folk psychology is "a framework of concepts, roughly adequate to the demands of everyday life, with which the humble adept comprehends, *explains, predicts* and manipulates a certain domain of phenomena" (Churchland, 1989, p. 225, emphasis added). Additionally, even as he argues against the place of folk psychology in a naturalistic account of human behaviour, arguing that it must be eliminated from our view of behaviour just as phlogiston was eliminated from our view of combustion, he recognizes that whatever else we do with folk psychology we use it "successfully to *predict* the future behaviour of others" (Churchland, 1989, p. 228, emphasis added) with "a facility…that is remarkable" (Churchland, 1981, p.68).

These views are also echoed by Churchland's primary opponent in the debate, Jerry Fodor. For example, he writes that "[t]he point – to repeat – is that the theory from which we get

this extraordinary *predictive* power is just good old commonsense[17] belief/desire psychology." (Fodor, 1987, p.3, emphasis and footnote added) Here we see a clear commitment to both the purpose of folk psychology – to predict behaviour – and the components of those predictions – beliefs and desires. Additionally, Fodor adds that folk psychology helps us to "infer people's behaviour from their intentions" (Fodor, 1987, p.3), which I take to be another way of expressing the explanatory power of folk psychology.[18] Similarly, in another paper Fodor (1991) immediately refers to "commonsense psychological explanation" and "commonsense intentional explanation/prediction" (Fodor, 1991, p.19) expressing a clear commitment to the idea that folk psychology really is an explanatory and predictive enterprise.

It is worth noting that Churchland and Fodor sit on one side of the debate regarding how to best conceive of our folk psychological practices. In particular, they both subscribe to the view that folk psychology is a theory; a view that has become known as Theory-Theory (TT). Briefly, theory theorists think that folk psychology is a term-introducing theory comprised of unobservable mental states and a set of principles, rules, laws, or generalizations that link these mental states together to generate behaviour. Notably, and most importantly, theory theorists think that the structure of folk psychology is very similar in nature to scientific theories (a point we will return to shortly). In particular, they think that the laws or generalizations that link mental states together to produce behaviour have the same structure as other scientific theories. This is made obvious by Churchland when he states that folk psychology "is so *obviously* a theory that it must be held a major mystery why it has taken until the last half of the twentieth

[17] Commonsense psychology is simply just another term for folk psychology.
[18] It is worth noting that early theorists in this debate often portrayed explanation and prediction as being symmetrical, as such, it is not unusual to read these two terms being used interchangeably or for these philosophers to focus on one of the two terms. See Andrews (2003), who identifies and then rejects the thesis that folk psychological explanation and prediction are symmetrical.

century for philosophers to realize it." (Churchland, 1981, p.71) On this view, the folk, as explainers and predictors, postulate the existence of some unobservable states within a person and calculate according to a set of rules until a behaviour is produced from the process.

One of the more notable and early alternatives to this view, called Simulation-Theory (ST), views our folk psychological practice not as one that unfolds on the basis of a tacit theory that we use and apply to ourselves and others, but rather, as a process of putting ourselves in the shoes of others and imagining what we would do if we were them. In other words, we explain and predict behaviour by taking mental states that we know others to have and running them through a simulation of what we would do or imagining what mental states we would have had if we engaged in the same behaviour as they did. In the words of Nicholas Humphrey (1986), "[w]e could…imagine what it's like to be [others], because we know what it's like to be ourselves…[I] make sense of [other's] behavior by projecting what I know about *my* mind into *them*." (pp. 71-72) As such, for simulation theorists the goal is not to employ some theoretical third person framework that dictates how and when mental states apply, but rather, to infer from one's own experiences what others in similar situations would believe, desire, intend etc.

Importantly, however, proponents of this view are aligned with Churchland, Fodor, and other theory theorists in a crucial way. They are still very much committed to the idea that folk psychology is an explanatory and predictive practice; that the goal of folk psychological practices is to explain and predict behaviour. In some sense, this must be the case if ST is going to be a genuine alternative to TT. In other words, ST was developed as an alternative way of explaining how it is that we engage in folk psychological practices, and there is a sense in which there ought to be some alignment between the two views in terms of their characterization of the purpose of folk psychology if they are to be true alternatives. The debate would be quite different

if ST and TT did not share this common starting point.[19] Either way, there is good evidence from key ST theorists that they do in fact share the same view as their TT counterparts.

Consider Robert Gordon (1986), who introduces his seminal "Folk Psychology as Simulation" with a discussion of the accurate predictions he has recently made using nothing but his folk psychology (p.158) and later when he writes "in *explaining* one's own behaviour, it would seem that one can – without invoking or using laws or theories – simulate *in imagination* various counterfactual conditions" (Gordon, 1995, p.115, emphasis added in first instance). Here we again see a clear commitment to the idea that the purpose of folk psychology is to facilitate the prediction and explanation of behaviour, the only change for Gordon is that this process is made possible by simulative practices that we undertake when we attribute beliefs and desires to one another. It is also worth noting that his seminal paper explicitly introduces and provides a briefly defended "new account of belief attribution" (Gordon, 1986, p. 158), committing Gordon as well to the claim that the attribution of propositional attitudes is the core activity of our folk psychological practices. Alvin Goldman, a key proponent of ST, also commits himself to the view that the purpose of folk psychology is to explain and predict behaviour. He writes, "*[p]redicting* another's decision is a stock example that ST aims to explain" (Goldman, 2006, p.19, emphasis added) and that "*explanation* can consist of telling a story that eliminates various alternative hypotheses about how the event in question came about or could have come about" (Goldman, 1995, p.89, emphasis added). Lastly, commitment to the view that folk psychology has the purpose of generating predictions and explanations of behaviours among simulation theorists is captured perfectly when Stich and Ravenscroft (1996) write that "simulation-theory

[19] Much of the debate is really premised on the idea that TT and ST are precisely two alternative accounts of the same thing. If ST theorists did not share the same underlying view that folk psychology is an explanatory and predictive enterprise, then in the strict scientific sense they are not so much alternative theories as they are competing paradigms or ways of viewing the practice itself.

provides a real alternative to the prevailing explanatory strategy in cognitive science [i.e. the theory theory] for explaining our capacity to *predict and explain* other people's behavior." (p.133, emphasis added) Here there is no question that Stich and Ravenscroft view their account of the mechanisms that underlie our folk psychological practices as an account of how it is that we predict and explain behaviour.

While these quotations should make it clear that there is a view that construes folk psychology as an explanatory and predictive enterprise that uses the attribution of mental states to comprise these explanations and predictions, the acceptance of this assumption and the role propositional attitudes play in folk psychology becomes exceptionally clear when we examine introductory textbooks on the philosophy of mind. Consider the following quotations:

> Much of our ordinary psychological thinking and theorizing ("commonsense" or "folk" psychology) involves propositional attitudes; we make use of them all the time to explain and predict what people will do. (Kim, 2006 p.15)
>
> 'Folk psychology is a conceptual framework' and/or 'network of principles' (perhaps largely implicit) used by the folk to understand, explain, predict their own and other's behaviour and mental states. (von Eckardt, 1994, p.300)
>
> The practice of explaining behavior by reference to the propositional attitudes is sometimes labeled 'folk psychology'. (Heil,[20] 2004, p.152)
>
> [F]olk psychology is the tag given to ordinary talk about the mind. It does not refer to talk about the biology of the brain and central nervous system; rather it refers to talk about beliefs and desires, intentions and fears, wishes and hopes. It is essentially the vocabulary we use to talk about and explain ourselves and others. It is the vocabulary of the mental. (Christensen & Turner, 1993, p.xvi)

[20] It is worth noting, for the purposes of full disclosure, that Heil does not personally think folk psychology is committed to revealing the underlying mechanisms of our behaviour. Instead, he thinks it might be useful and perhaps even indispensable, but he is not committed to any stronger claim about folk psychology being causally accurate.

> We seem to be strongly committed, in our commonsense or 'folk psychological' ways of thinking about people, to the idea that attitudinal states with abstract propositional content can legitimately be invoked in causal explanations of people's actions. (Lowe,[21] 2000, p.78)

These quotations make it plainly clear; in the literature that has the sole purpose of educating a new generation of philosophers or capturing the history of philosophical thinking regarding the mind in an encyclopedia, there is a strong commitment to describing folk psychology as an explanatory and predictive enterprise. In each case it is emphasized that folk psychology is what we are using when we explain and predict behaviour.[22] The fact that, for the purposes of education and to bring others up to speed on the philosophical debate, folk psychology is being characterized as an explanatory and predictive framework, that involves the attribution of propositional attitudes, gives us good reason to think that these assumptions really do underlie the traditional construal of folk psychology.

It should be clear by this point that there is a prominent and traditional construal of folk psychology such that the primary purpose of folk psychology is to explain and predict behavior. This is held by a range of philosophers each with various and often competing commitments regarding how it is precisely that we achieve this goal. Additionally, this view is espoused by those providing summaries of the debate for other philosophers or for the future generation of philosophers. As I noted above, the idea that the primary purpose of our folk psychological practices is to explain and predict behaviour does sound plausible. To repeat myself, what could

[21] Lowe is also a bit more skeptical about the nature of folk psychology. In particular in capturing his own view he writes "[b]ut perhaps we should view with a more skeptical eye the comfortable assumption that beliefs and desires are straightforwardly 'states' of persons, or of their minds, by analogy with such physical states of their bodies as shapes, mass and velocity." (Lowe, 2000, p. 297) I think this is important, as he recognizes that the tradition is as he's described, but questions whether or not propositional attitudes are really analogous to scientific entities.

[22] I should emphasize that when I say there is a commitment to these claims, I mean merely that there is a commitment to the claim that folk psychology is traditionally understood in this way. That is, I am not committing Heil, Kim etc. to the claim that folk psychology is this way, instead I am simply committing these philosophers to the claim that folk psychology is traditionally understood in this way.

be a better way to navigate the social domain than to provide explanations and predictions of ourselves and those around us? The intuitiveness of this position alone may be enough to justify how it is that this view became so commonplace. But there is another more theoretical and speculative argument for thinking that the purpose of folk psychology is to explain and predict behaviour that I want to turn my attention to in order to explore the plausibility of this view a little more and to demonstrate how it is that this view came into vogue.

2.4 – Speculative Anthropology

Support for the view that folk psychology is an explanatory and predictive enterprise stems also from the claim that explanation and prediction would be, evolutionarily speaking, advantageous and could be selected for on this basis. While we do not know for sure exactly when humans started to attribute mental states to one another (and ourselves) we can follow Wilfred Sellars and engage in some speculative anthropology about how this skill may have evolved and how it relates to navigating the social domain through repeated explanation and prediction.

In his famous *Empiricism and the Philosophy of Mind*, Sellars (1956/1963) invites us to consider a myth, or more precisely, an anthropological science fiction story. The cast of the story is the human race at some point in history before mentalistic discourse is the norm. Importantly, the human race at this stage in history is still able to reason quite effectively and have developed a primitive but sophisticated language that is Rylean[23] in nature. That is, while the language contains all the logical expressions (i.e. conjunction, disjunction, negation and quantification) it deals only with public objects. Thus, with respect to other humans the language only addresses overt behaviour, be it verbal or physical, and makes no reference to any object or item that is not in the public space and shared with others. Sellars begins with this assumption for a reason.

[23] Of course Gilbert Ryle (1949) is famous for arguing that locutions describing mental states are conceptually equivalent to ones describing behavioural dispositions.

Specifically, he does not find it surprising or puzzling that humans developed a language to refer to public properties or objects (Sellars, 1956/1963, p.178). In fact, his discussion takes it for granted that a Rylean language is an uncontroversial starting point. What is puzzling to Sellars, however, is how it is that we learned to speak of inner episodes and immediate experiences (Sellars, 1956/1963, p.178).

In other words, how did our Rylean ancestors evolve to be like us; humans with a full mentalistic discourse that permits us to effortlessly speak of the internal unobservable properties of the mind? In Sellars' own words "[w]hat resources would have to be added to the Rylean language of these talking animals in order that they might come to recognize each other and themselves as animals that *think, observe,* and have *feelings* and *sensations,* as we use these terms?" (Sellars, 1956/1963, p. 179, original emphasis) and perhaps more importantly, "[h]ow could the addition of these resources be construed as reasonable?" (Sellars, 1956/1963, p.179) In other words, there must be some evolutionary explanation for how it is that we came to speak about the mental, going beyond our purely public discourse.

Sellars invites us to consider Jones, a Rylean ancestor who just so happens to be a genius. Jones operates quite successfully with his Rylean language, observing his fellow humans going about their business acting and speaking out loud in conjunction with this activity. One day, however, Jones notices that not all intelligent looking behaviours are accompanied by overt verbal behaviour. That is, while often behaviour is accompanied by what we now characterize as "thinking out loud" behaviour, this is not always the case. One day Jones might observe Smith engaging in some behaviour accompanied by his typical verbal expressions, but the next day Jones observes the same behaviour but without the verbal expression. This observation, in

conjunction with others, leads Jones to conclude that not all intelligent behaviour must be accompanied by the typical 'thinking out loud' verbal expressions.

Jones is now faced with a gap in his knowledge. His Rylean language and the understanding of others that it imparts on him permits him to only understand behaviours that are publicly shared, but now a behaviour that he would have attributed intelligence to no longer fits within the confines of his limited language. In order to come to grips with this new phenomenon Jones must go beyond his Rylean language and come to understand how people can act intelligently when their behaviour is not "threaded on a string of overt verbal episodes" (Sellars, 1956/1963, p.186).

In order to account for this new explanatory gap, Jones develops a theory that at its core assumes that something like the overt verbal utterances he is used to observing must be occurring internally along with those intelligent behaviours. There must be some internal process that accompanies the behaviour that he cannot witness, but that causes the behaviour. To put it at its crudest, Jones can assume that the typical utterance that happens overtly happens internally as well; that there is an internal speech that goes along with the silently intelligent behaviours.

However, once this step is made it is not a big step to realize that the outward speech itself might really just be the culmination of this internal process and that the outward speech does not cause the action, but instead that the internal process that is there to cause the behaviour in the silent cases is there in the noisy ones as well. This inner speech is the root cause of the behaviour in both the silent and expressive cases. Thus, Sellars concludes "even when a hungry person overtly says, 'Here is an edible object', and proceeds to eat it, the true-theoretical-cause of his eating, given his hunger, is not the overt utterance, but the 'inner utterance of this sentence'." (Sellars, 1956/1963, pp.186-7)

In effect, Jones has developed a primitive theory of human action that allows him to understand, explain, and predict behaviour in a way that is more expansive than his Rylean language permits. While it is true that this theory requires that he attribute unobserved, non-empirical, inner episodes to those around him (Sellars, 1956/1963, p.187) he has good reason to engage in this theoretical practice giving his theory a "methodological purity" (Sellars, 1956/1963, p.187). This attribution could also be viewed as giving Jones and those who he shares this learned skill with, a competitive advantage over others. The advantage imparted on Jones through the development of this new skill helps understand both the introduction of the skill as being reasonable and the persistence and refinement of this skill over time.

Ultimately and most importantly, what we have then is the beginnings of a theory that recognizes humans as thinking things with internal states that are attributable on the basis of their relation to observable behaviours. Importantly, this theory can be taught to others and once mastered can be applied to oneself. What we see then is the beginnings of attributive practices that can be refined and perfected over time eventually becoming what we now recognize as folk psychology. The act of thinking can be broken down into categories and specific theoretical relationships can be defined. In effect, once the theory is constructed, we can continue to build it adding new components or objects and refining the properties of them. The result is a theory that allows us to attribute inner states to ourselves and those around us permitting us to develop explanations and predictions of behaviour.

What is important about Sellars' tale is that it is a very early account of folk psychology as a practice used by the lay person that underpins the traditional construal explored above,[24] but that gives us a plausible answer to the question of "why do we engage in folk psychological

[24] As noted below, particularly the TT incarnation of the traditional view.

practices?" The answer is, that the explanation and prediction of behaviour using the folk psychological or mentalistic discourse provides us with a more comprehensive understanding of those around us that ultimately improves our ability to explain and predict behaviour.

It is worth noting that Sellars is an early proponent of TT and the tale he tells is certainly one that describes folk psychology as a term-introducing theory. However, I don't believe this piece of the story is necessarily crucial to the more general point he advocates for, more specifically, the general evolutionary story that he provides. Whether folk psychology is ultimately theory based, simulation based, or some hybrid of these two views, so long as it operates as an explanatory and predictive enterprise the speculative anthropological or evolutionary tale Sellars provides can be picked up and adopted to support the view.[25] Importantly, however, Sellars is not the only philosopher to claim that the explanatory and predictive nature of folk psychology imparts on it some evolutionary advantage.

George Graham (1987), for example, re-emphasized the suggestion put forward by Sellars, arguing that the "adaptive biological utility" of folk psychology as an explanatory and, in particular, predictive enterprise "helped us to survive and proliferate" (p.358). It is worth noting that Graham ultimately defends and adopts a TT version of folk psychology using the term "concept of mind". In particular, through his analysis of the adaptive biological utility of folk psychology he effectively analyses and defends the "concept of mind" view of folk psychology arguing that it best aligns with the adaptive story.

[25] We might, for example, suggest that instead of Jones recognizing that the verbal utterances that normally drive someone else's behaviour are not always observable and so Jones theorizing that there are internal states that are really causing the behaviour, Jones could recognize that his own behaviour is caused even without verbal utterances and learn that his own 'internal' dialogue is the source of his action. From this personal experience he could begin to explain and predict the behaviour of others by asking "what would I do in their situation?", thereby utilizing a simulation in order to predict the behaviour. In other words, Sellars' story may be easily adaptable to the ST account of folk psychology.

In contrast, however, Nichols and Stich (2003) who defend a hybrid view of folk psychology, one that is part simulation part theory, make a similar move by developing their own "Just So Story" to account for the plausibility of their own view (pp.60-68). While some might view these speculative stories as mere fictions,[26] Nichols and Stich note they are theoretically useful devices. In particular, they note that "[i]f there is nothing in the story that conflicts with the known biological and psychological facts or with what little is known about the actual origins of the mechanisms in question, then the story provides an account of how the mechanisms we posit might *possibly* have evolved." (Nichols & Stich, 2003, p.60) and they ward off any objections regarding whether or not the posited view makes biological sense. Most importantly for our purposes here, Nichols and Stich focus their "Just So Story" squarely on the predictive success of folk psychology arguing that this success makes sense of how it is that the practice evolved and was selected for. In fact, from their perspective without this predictive success it might be a mystery how we developed our folk psychological competence since "[f]or as far as Mother Nature is concerned, that [prediction] is how mindreading[27] first earned its keep." (Nichols and Sitch, 2003, p.67, footnote added) The point is clear, for Graham, Nichols and Stich, arguably Sellars, and others,[28] the explanatory and predictive success of folk psychology is precisely the reason it was selected for and persists. It is this type of argument, I think, that helps us to make sense of how it is that the traditional view explored above has come to be so widely and prominently held since it gives us some reason to think that folk psychology ought to be this way. However, this discussion also identifies a slew of additional philosophers who view folk psychology as an explanatory and predictive enterprise. At this point, we have good reason to

[26] A story I have heard on occasion when I raise the plausibility of TT on the basis of Sellars' story.
[27] The term "mindreading" is synonymous with the attribution of propositional attitudes to others.
[28] See for example, Baron-Cohen, 1995; Goldman, 2006.

think that traditionally construed folk psychology is viewed as an explanatory and predictive enterprise. What remains to be demonstrated is the related claim, that traditionally construed folk psychological explanation and prediction are understood in at least quasi-scientific terms.

2.5 – Folk Psychology and Science

The claim that folk psychological explanation and prediction are to be understood as being science-like is not always explicitly identified in the literature, although sometimes it is and we'll see this below. But it is worth looking at the intuitiveness of idea first, and what it really means.

One way of looking at this idea is to say that folk psychology and science share some basic commitments or characteristics. Of course, if folk psychology is like science, it is likely that the similarities are somewhat rough in nature. For example, it is all too likely that science is "more rigorous, more systematic and more explicit" (Knobe, 2006a, p.7) than we could ever expect folk psychology to be. Of course, this increased rigor, the systematic nature of science and the fact that science is more explicit in its categories, goals, and domain of applicability does not mean that folk psychology cannot share in these characteristics. Perhaps then, we can look at folk psychology and science as being two points on a shared spectrum. This idea is also espoused by Sellars who sees common sense and scientific inquiry as being continuous – not distinct enterprises that differ in kind, but rather enterprises that differ in degree in terms of the applicability or expression of some characteristics (Sellars, 1956/1963, pp.182-183).

Speaking broadly about the folk and the scientist, Sellars writes "the way in which the scientist seeks to explain empirical phenomena are refinements of the ways in which plain men, however crudely and schematically, have attempted to understand their environment and their fellow men since the dawn of intelligence." (Sellars, 1956/1963, p.183) This general analysis of the similarities between common sense and scientific approaches to understanding the world

around us can be applied to the domain of human behaviour and so folk psychology can and ought to be viewed as science-like on Sellars' view of the relationship between common sense and scientific understandings. Whether we think that there is a loose overlap between folk psychology and science, that they are two enterprises existing at different points on the same continuum, or that folk psychology and science are aligned in all the right ways, there is a general characteristic that is being described in all these possible relationships that I will capture using the term "quasi-scientific". This term allows for some flexibility in terms of the degree to which there is alignment between folk psychology and science, but captures the general sentiment that there are indeed some similarities.

I think the term "folk psychology" itself also identifies the expected relationship between this practice and science. 'Psychology' just is the scientific study of the mental and of behaviour and to include the name of this practice in the term "folk psychology" instantly creates a perceived relationship between folk psychology and science. However, this relationship is then mediated by the application of the adjective "folk" to the practice of "psychology". Obviously this is meant to distinguish *folk* psychology from psychology *proper*, but retain the core elements of psychology. While this etymological analysis is not exactly proof of the claim, it does capture the intuitiveness of the association between folk psychology and science.

The idea that folk psychology is quasi-scientific is also evidenced, in part, by the commitment to explanation and prediction itself. With explanation and prediction as the goal of the folk psychological practice, there is a sense in which folk psychology is immediately aligned with science. That is, more often than not the goal of science *just is* the explanation and prediction of some natural phenomenon. As such, pitching folk psychology as an explanatory and predictive enterprise immediately aligns this practice with the general practice of science.

While there are potentially other ways of construing explanation and prediction, construing them as scientific is the most obvious and intuitive assessment and I believe that most philosophers describing folk psychology in this way have been doing so with the assumption that we all just intuitively use these terms in the same way. This is not to say that the methods of scientific inquiry and folk psychology are necessarily the same, although they most certainly could be (see comments above), just to observe that the goals are the same. Further support comes from the observation that not only do folk psychology and science share the same goal (e.g. explanation and prediction) but that they can have and in fact do share a target of study. More specifically, both folk psychology and science are used to understand social and mindful behaviour. This too adds to the intuitiveness of the alignment between folk psychology and scientific inquiry.

Often a scientific inquiry is assessed on the basis of its predictive and/or explanatory utility. Similarly, scientific inquiry is often very domain specific and is often used as a special-purpose tool (Knobe, 2006a, p.13). As so far explored, folk psychology is also very much a domain specific tool – applying only to the domain of social and mindful behaviour. This is further constrained by the characterization that folk psychology allows us to attribute mental states to ourselves and others in order to explain and predict behaviour. As such, the view of folk psychology explored thus far is very much constrained in the same kinds of ways that scientific inquiry is constrained and the specialization of folk psychology is similarly quite significant.

Likewise, as folk psychology is a tool for just explaining and predicting behaviour it has been assessed on this basis, and assessed quite favourably. In fact, in the discussion above we saw a clear focus on the predictive success of folk psychology as being evidence for its adaptive utility, its creation, and its sustained existence. We even saw philosophers on both sides of the

realism debate regarding propositional attitudes, namely Churchland[29] and Fodor, speaking to the success of folk psychology as a predictive and/or explanatory enterprise. This too reveals that folk psychology, traditionally construed, has been evaluated in the same way that we evaluate scientific enterprises. This, reasonably, leads one to conclude that folk psychology is at least in some way, a quasi-scientific enterprise.

All of this discussion suggests that it is intuitive to think of folk psychology in the same general terms that we think of scientific inquiries. The intuitiveness of this relationship is so obvious to some in the profession that to even articulate it seems unnecessary and to challenge it seems strange and raises many eyebrows. In fact, when proposing a suggestion to the contrary, myself and others[30] often receive significant resistance.[31] Fortunately, we need not rely just on this intuitiveness or the resistance myself and a few others experience when presenting the alternative view. Instead, some philosophers involved in the debate over folk psychology have clearly expressed their commitment to the idea that folk psychology is quasi-scientific. This is most obvious for those who view folk psychology in theoretical terms.

As soon as we hear the word 'theory' there is an immediate association with science that cannot be helped; the intuitiveness of this relationship is almost unavoidable. In fact, the TT approach to understanding our folk psychological competence has not been particularly secretive

[29] While Churchland ultimately argues for the elimination of folk psychology, as discussed above, he does recognize that folk psychology is a reasonable tool for the humble demands of everyday life (Churchland, 1989, p. 225) and that it does permit us to successfully predict the future behaviour of others with a fairly remarkable degree of success (see Churchland, 1981, p. 68; Churchland, 1989, p. 228).

[30] See for example Knobe (2006a) who discusses a very similar experience when making a related claim. In particular, he notes the "incredulous stares" (Knobe, 2006a, p. 16) he received from others when claiming that folk psychology has a moral character to it that distinguishes it significantly from scientific inquiry.

[31] Most notably, a prominent philosopher once told a room that I was obviously wrong, because folk psychology *just is* scientific and that beliefs and desires *just do* cause behaviour. He was literally shocked that I could claim anything to the contrary.

about the idea that folk psychology is science like. In fact, simply characterizing folk psychology as a term-introducing theory immediately associates it with science.

Churchland (1981) goes to great lengths to demonstrate that folk psychology is a theory in exactly the scientific sense. Churchland first notes that the explanations provided by folk psychology presuppose that there are laws connecting "the explanatory conditions with the behaviour explained" (Churchland, 1981, p.68) and then argues that these laws – while "rough and ready" – have the general structure of other scientific laws. Moreover, he notes that propositional attitudes (e.g. …believes that *p*) and what he calls the "numerical attitudes" of the physical sciences (e.g. …has a mass_{kg} of *n*) have exactly the same structure and figure into the laws of the theory in the same kind of way (Churchland, 1981, p.70). In fact, Churchland's use of the term 'theory' to describe folk psychology cannot be interpreted in any other way other than as a scientific theory.

Similarly, the whole debate that takes place regarding the status of folk psychology, the realism debate regarding folk psychology, is essentially premised on the alignment of folk psychology and science. For example, Churchland conceives of folk psychology as a theory which allows him to ask whether or not folk psychology is true and then to argue for the replacement of folk psychology by another scientific framework – neuroscience (1981) or connectionist theory (1989). Without the alignment of folk psychology with science, this argument would not get off the ground.[32]

Similarly and not surprisingly, Fodor also conceives of folk psychology as an explanatory and predictive enterprise that *just is* aligned with science. In fact, he writes "commonsense belief/desire psychology explains vastly more of the facts about behaviour than any of the

[32] I will discuss claim in more detail in Section 6.5 of Chapter Six.

alternate theories available" (Fodor, 1987, p.x) where the alternative theories he has in mind are alternate scientific theories. Moreover, and more importantly, Fodor conceives of folk psychology as a scientific theory that will be vindicated by the representational theory of mind, effectively a computational psychology. Again, he assumes that the relationship between folk psychology and science is such that the categories and laws of folk psychology will be appropriately vindicated by science. As von Eckardt (1994) notes in her analysis of the realism debate, both sides effectively "share the assumption that the ultimate arbiter of the reality of [folk psychology] and the trust of [folk psychology] generalizations *is* science." (p.305, emphasis added) In this way, folk psychology is portrayed as being appropriately aligned with science such that we can evaluate alternatives and assess the truth or falsity of the explanatory and predictive enterprise.

The idea that folk psychology is a scientific endeavour is also picked up by those in developmental psychology. I noted above that Sellars views common sense and science as existing on a continuum and a similar idea is endorsed by the developmental psychologist Alison Gopnik and her colleagues. In effect, they characterize children's use of folk psychology as operating like a *child scientist*[33] who comes to possess their "theory of mind" or folk psychology over time. What is most telling about their characterization of children is that they argue based on the evidence of their studies[34] that children revise their theory of mind in the same ways that scientists revise their theories. That is, that new evidence is collected and this new evidence is used to revise a theory (see for example, Gopnik & Meltzoff, 1997; Gopnik & Wellman, 1992).

[33] It's worth noting that this view is often referred to as the child scientist theory.
[34] Which of course there is significant debate about. The purpose of identifying this view is not to endorse it, but rather to support the claim that folk psychology is often portrayed as being aligned with science.

By focusing in on the development of folk psychology[35] among children, Gopnik and her colleagues have effectively claimed to observe Sellars' genius in action. They are watching children develop their theory in the same way scientists do, leaving adults with a fully functioning folk psychological competence as the product of a long and hard fought scientific endeavour.

Whether simulation theorists are committed to the same alignment between folk psychology and science is much less clear. Obviously though, insofar as simulation theorists think of ST as being an explanatory and predictive enterprise, there is alignment between folk psychology and science just in the way I explained above; albeit a potentially weak association. That is, there is an alignment of the goals and the object of interest between the two enterprises. However, insofar as simulation theorists reject the idea that folk psychology requires a theoretical competence the obviousness of the relationship of their version of folk psychology and science is less transparent.

That said, *pace* Gordon, ST theorists have started to allow for or recognize that their version of folk psychology is incomplete without some theoretical competence to complement the simulation mechanism they describe.[36] These hybrid accounts are built on the principles of both ST and TT. Most prominently, Nichols and Stich (2003) and Goldman (2006) argue for a particular brand of ST that carves out a space in which theory must play a role in our folk psychological competence. Importantly for my purposes here, to the extent that simulation

[35] Of course this view of children as proto-scientists extends beyond just the folk psychological realm, but I focus here just on this narrow application of the child scientist view.

[36] For example, the simulation story still unfolds, but we need to have some bare recognition that folk psychology is a term-introducing framework that identifies unobservable entities. The difference between TT and ST is in the particular way in which we deploy these terms. TT deploys them in an objective third person way, while ST relies on the simulation mechanism to attribute them to others.

theorists grant a role for theory in their analysis of folk psychology, the relationship between folk psychology and science will be more robust as it is in the analysis I presented above for TT.

There is, however, another way in which we can analyze the claim that folk psychology is quasi-scientific. Thus far I've looked at the relationship between folk psychology and science in terms of the general goals and structure. But we can go one step further as well, and assess the extent to which traditional theorists are committed to the claim that folk psychological explanations and predictions purport to pick out the causes of behaviours.[37] To the extent that the traditional conception of folk psychology is characterized in such a way that it purports to pick out the actual causes of our behaviour, there is an obvious alignment with science yet again. Naturally, the goal of science is to explain and predict some natural phenomenon by identifying the causal structure underlying what it is that we are observing.[38] If the beliefs and desires that are cited in our folk psychological explanations and predictions are meant to be the actual causes of the behaviour we are trying to explain or predict, than there is even more reason to think that folk psychology is being construed in quasi-scientific terms. Again, the story is clear and transparent for theory theorists but a little less transparent for simulation theorists.

The TT characterization of folk psychology just simply is a causal analysis. More specifically, TT theorists believe that folk psychological explanations and predictions are generated through the application of a causal theory of behaviour. For example, Churchland

[37] We can again trace this suggestion back through the history of philosophy of action and again back to Aristotle. As noted in footnote 15 above, Aristotle's analysis of practical wisdom as a behaviour producing mechanism was a logical model of behaviour; propositions and moral prescriptions combined in logical ways to produce conclusions regarding action. This view was criticized as being causally inert and so we could, for example, follow Hume and build causality into the model of behaviour by combining both reason and desire. Importantly, this move never rejected the logical analysis of behaviour that was central to the position advanced by Hume. As such, we have seen a move in the history of the philosophy of action from the logical to the causal; preserving the logical structure of behaviour while building causality into our analysis of action.

[38] For example, chemical theory is predicated on the idea that it captures the causal interactions of various chemicals in order to predict or explain the resulting chemical reactions etc.

claims that “the bulk of the generalizations of folk psychology are causal or nomological” (Churchland, 1989, p.228) and that “the simpler parts of folk psychology are transparently causal or nomic” (Churchland, 1989, p.230). Fodor echoes this sentiment when he writes that one of the conditions of his view just is that mental states are causal (Fodor, 1987, pp.10, 12). Lastly, and as we saw above, Sellars echoes this sentiment in his speculative story noting that Jones comes to recognize that there are inner states causing the behaviour just as the overt verbal utterance does. This commitment makes sense when thinking about why it is that explanation and predictions of behaviour using folk psychology are viewed as being so successful. The suggestion that a prediction accurately picks out the causes of the thing being predicted helps to make sense of why the prediction is accurate. In the context of Jones the genius, his theory accurately captured the causes of behaviour which allowed him to explain and predict things in a way that imparts an evolutionary advantage.

The picture for ST theorists is a little less obvious and explicit. Again, to the extent that simulation theorists adopt theory into their view, they may be party to the same causal commitments that theory theorists have explicitly made. Additionally, any commitment to explanation and prediction as being the goal of folk psychology may also just implicitly endorse the idea that the explanation and prediction are correct because they pick out the causally relevant factors and arrange them in the appropriate way to provide an explanation or prediction. This suggestion integrates well into an observation about the ST account of folk psychology.

ST theorists are committed to the idea that within each of us there is some action control system (or similar term). This action control system does precisely what it sounds like it does, it controls our actions. Given the focus of the debate regarding folk psychology on propositional attitudes, the claim then is that our action control system produces behaviours depending on the

propositional attitudes we hold. In this sense, the propositional attitudes we hold and the action control system together cause our behaviour. The trick for simulation theorists is that when we are explaining or predicting the behaviour of someone else, we are just utilizing this action control system "offline" in order to run a simulation of what we would do or why we would have done what we did given the same circumstance the other is in. We must take our system offline because if we did not, holding these propositional attitudes in our action control system would actually cause the behaviour. However, when we attribute a belief or a desire to another person, we are effectively committing ourselves to the claim that they have a particular set of beliefs and desires that when fed into their action control system (which we assume to be appropriately similar to us) will cause them to behave in a particular way. In effect, we have made a prediction or an explanation on the basis of a causal story. In this way, even though ST is based on the idea of simulation or pretence, there is very good reason to think that a pure ST analysis of folk psychology would still very much be a causal model.[39]

Importantly, the idea that folk psychology is purported to pick out the causes of our behaviour is also captured by others. Returning to the quote from Lowe above which has been pulled from an introductory textbook to the philosophy of mind, we clearly see a characterization of folk psychology as a causal enterprise. Specifically, he explicitly identifies the assumption that "propositional attitudes that figure into folk psychological explanations are appropriately causal in nature" (Lowe, 2000, p.78).[40] Others have more recently made a similar claim about the traditional conception of folk psychology, in their own attempt to reject this traditional view. For example, Tadeus Zawidzki (2008) notes quite clearly that folk psychology has traditionally been

[39] I recognize that Gordon tries to explicitly reject the idea that his model is causal. It is unclear to me whether this is a successful move and it does not seem as though this would necessarily apply to all ST theorists.
[40] Again Lowe is skeptical of this view himself, but recognizes it as being an accurate description of folk psychology.

construed as having the task of describing the causes of behaviour in order to predict a person's behaviour (p.199). Construing folk psychology in this way is not particularly strange in the context of the history of philosophy. The discussion of folk psychology really emerged out of a history of philosophy of mind and philosophy of action, both of which are very much interested in the causes of behaviour or the causal nature of the mental and the brain. Philosophers have been interested in the causes of behaviour and I believe this has been projected on to their analysis of the folk as well.

In all of the above senses, I think that the traditional construal of folk psychology is very much committed to the idea that folk psychology is at least a quasi-scientific enterprise, where explanation and prediction are to be understood in loosely scientific terms and that accurately describing the state of affairs that caused or will cause a behaviour is the goal of the project. For some philosophers this will be transparently the case and for others we might have to do a bit more work. But generally speaking there is good reason to think that those interested in folk psychology traditionally have had a quasi-scientific and causal view of folk psychology in mind.

2.6 – The Traditional Construal of Folk Psychology

There are many other ways to construe folk psychology.[41] The purpose of this chapter has not been to argue that every philosopher who has ever thought about folk psychology has thought the same things and in particular, have conceived of folk psychology as an explanatory and predictive enterprise that is quasi-scientific in nature.[42] The point of this chapter has been to

[41] Most notably, Model Theory (e.g., Maibom, 2003), Rationality Theory (e.g., Dennett, 1987), Interactionist accounts of folk psychology (e.g. Gallagher, 2001), and even Narrative accounts of folk psychology (e.g. Hutto 2008).

[42] It is worth noting, however, that many of the philosophers identified in fn. 41 do in fact operate within the assumption that folk psychology is an explanatory and predictive enterprise. Where they differ is in their particular analysis of how folk psychology functions as an explanatory and predictive enterprise. Defending this claim for each of the above approaches to understanding folk psychology is simply beyond the scope of this particular project and would only serve to *reinforce* the point already argued for and sufficiently demonstrated above. However, I acknowledge that this is a worthy project and one I would happily engage in, but that I cannot engage in this

identify some core commitments that are generally held by philosophers interested in folk psychology; a set of commitments that are both explicitly or implicitly endorsed, and potentially endorsed to varying degrees. More specifically, and as noted throughout this chapter, the traditionalist is committed to the idea that folk psychology is a practice with the goal of explaining and predicting behaviour, and that these practices should be understood in, at least, quasi-scientific terms. I've tried to identify some common ground within a fierce debate about the practice and while I think each of these assumptions are often held together by a number of philosophers, I grant that some philosophers may hold only some of these assumptions.

In the chapters that are to follow we will see that philosophers often hold assumptions about the folk that appear not to be true. In particular, we will see that the assumptions identified above are not true of the folk when they engage in folk psychological practices. Whether any philosopher holds one or all of the above assumptions, their project will be undermined to the extent that the empirical evidence we are about to explore has merit. The point of this chapter has been just to identify a few key assumptions that have underscored much of the traditional discussion regarding folk psychology in order to challenge them in what follows.

particular project. The goal here has merely been to set the stage with enough detail to motivate the idea that these assumptions are held by philosophers working in folk psychology. The project needs to progress to the next step at some point, and so I have resisted continuing to analyze each interpretation of folk psychology in full detail.

Chapter Three: The Experimental Method

3.1 – Introduction

In the last chapter I demonstrated that a prominent and traditional construal of folk psychology is committed to the claim that when the folk navigate the social domain, they do so by making quasi-scientific explanations and predictions about how people will behave and act. Unlike most works of philosophy, which at this point in the project would move towards developing arguments for or against the various positions covered in the previous chapter, this project requires an intermediary chapter much more akin to the "methodology" chapters found in social science research projects.[43] In this chapter I will detail the lens through which I intend to examine the assumptions that underlie the prominent and traditional construal of folk psychology.

Without intending any offence, philosophy is at least sometimes portrayed or viewed as a theoretical enterprise detached from empirical inquiry. Whenever I disclose my area of research to someone, they often make references to imagery of me sitting in an Ivory Tower or having my "head in the clouds". Even among other academics I hear comments that suggest we philosophers only *think* about things, never truly *studying* or *engaging* the actual, ugly, dirty, difficult world.[44] While these comments are certainly anecdotal in nature, I do not think they are unusual, nor are they comments that have only been directed at me. Moreover, it is not particularly unusual to come across pieces of philosophy that do not make mention of or

[43] This is not to say that all works of philosophy don't have a similar intermediary chapter. In some instances philosophers will identify a particular lens through which they want to examine a well-entrenched debate or issue. That said, the structure of this project is much more akin to a social science research project, with the following structure: Introduction, Background, Methodology, Results, Discussion.

[44] This is a comment that I've heard when working with or interacting with academics in the social sciences. The general thrust of the comment is that social sciences (and indeed most sciences) directly engage with the world in a way that most theoretical endeavours (philosophy included) do not. The point is simply that the world is really a challenging thing to study and so our approach to understanding it must similarly be challenging and perhaps imperfect as it tries to tackle and deal with the real world. Of course this is not to say there is not a role for theoretical analyses in this approach, just that we must get our hands dirty if we're going to encourage the growth of knowledge.

reference to any empirical research. This is not to say, however, that philosophy is never informed by our experiences, and entire approaches to conducting philosophical investigations have emerged from this particular commitment,[45] and it is not to say that philosophers are not often informed by empirical research, it is just to note that there is often a significant detachment between philosophy and any other empirical area of inquiry.[46]

Perhaps more importantly than the casual observations I've just made, there is the growing number of researchers who identify themselves as philosophers and yet paint a very similar conception of philosophy, one that is very detached from the empirical approach they adopt. These new[47] philosophers, grouped together originally under the image of a flaming armchair and the banner "experimental philosophy", have a particular brand of empiricism in mind, an experimental one. More specifically, the experimental method[48] is adapted to the philosophical setting in order to explore philosophical concepts and inform arguments regarding which positions philosophers can and should hold.

Notwithstanding all of the above comments, it would be unfair to make the wholesale claim that philosophers do not take into consideration empirical and/or experimental evidence.

---

[45] In the history of philosophy there is a clear distinction between two broad approaches to conducting philosophical investigations; "Empiricism" and "Rationalism". The former is meant to take seriously our experiences and use this information to inform our philosophical thinking, whereas the latter is meant to be an entirely theoretical enterprise that rationally deduces truths from axioms or other incontestable propositions that are revealed to us through intuition alone. The distinction I'm making here is narrower and focuses on the fact that philosophy is oftentimes detached from the empirical undertakings of science or social sciences.

[46] Whether philosophy should always be informed by our experiences is a question well beyond the scope of this particular project. Most obviously, the relationship between philosophy and empirical inquiry likely needs to be assessed on a case-by-case basis. In fact, as we'll see in the discussion below, the relevance of the particular brand of empiricism that I am advocating for varies depending on the philosophical issue we're exploring. The point here has only been to identify that there is often this detachment since this observation is relevant to the history of experimental philosophy.

[47] "New" is a relative term in a number of ways. The "new" approach to philosophy that I will recount in this chapter really got going in the early 2000's, which by this point in time might already make this approach somewhat old. However, there is also some reason to think that the principles of this approach as employed by some harks back to ancient skeptics who cast doubt on the arguments of anyone who moves beyond the appearances to theorize about the world that cannot be accessed directly.

[48] Using experiments to test whether or not a particular hypothesis is correct.

For example, those studying philosophy of physics need to be and are concerned with any experiments testing the latest theories in physics,[49] those considering the relationship between psychopathy and ethics need to consider empirical evidence regarding the nature of psychopathy,[50] ethicists are now becoming interested in neuroscience and a new sub-domain of ethics called "neuroethics" has been recently developed,[51] and by no means does the list end here. In fact, and this will be made more clear towards the end of the chapter, philosophers of mind have often been interested in empirical findings, including those pertaining to neurological disorders[52] or developments in cognitive science and artificial intelligence.[53] But in the chapters that follow, I will place a significant emphasis on the value and importance of empirical evidence as it relates to our understanding of folk psychology and in particular, I will recount experiments that I and other experimental philosophers have conducted to specifically test philosophical hypotheses regarding folk psychology. More specifically, I will put the assumptions that underlie the prominent and traditional account of folk psychology just explored to the test and ask whether experimental research validates or rejects these claims.

All of what follows may appear to have a radical quality[54] to it that warrants acknowledgment and discussion. As such, this chapter sets out to articulate the appropriateness of evaluating the traditional construal of folk psychology through an empirical lens. The use of

[49] See for example Ladyman & Ross (2007) who defend their view that the world has an objective modal structure on the basis of current physical theory and reconcile their view with the various approaches of the special sciences.
[50] See for example Maibom (2005) who uses empirical evidence regarding psychopathy to develop an analysis of competing moral theories.
[51] See Roskies (2002) for an early overview of the field and the questions that are trying to be answered by these researchers. Most importantly, an explication of the interest of researchers in understanding moral cognition, how values are represented in the brain, and how ethical decision-making unfolds.
[52] See for example Bayne (2011) where the implications of absence seizures are examined to inform our understanding of consciousness or Bayne (2008) where research on split-brain syndrome is evaluated in the context of the theory that consciousness is unified.
[53] See for example Eliasmith (2003) who argues for a representational and dynamical theory of cognition that is squarely rooted in empirical work regarding the brain, connectionism and cognitive modeling more generally.
[54] At least it is often portrayed as being radical and I grant that at minimum it is unusual for philosophy to take such an empirical bent and rely so extensively on the experimental method.

the empirical lens alone,[55] I suspect, will be fairly non-controversial and so I will address that towards the end of this chapter (see section 3.5 – The Empirical Lens). In contrast, the association of this project with experimental philosophy warrants a more comprehensive discussion for two reasons.

First, experimental philosophy still has a "new car smell" associated with it and so a full discussion of this approach to conducting or, as it is often portrayed, criticizing philosophy is warranted in the interest of full disclosure. Second, and more importantly, while a significant portion of the empirical evidence I will utilize was produced under the guise of experimental philosophy, I think there are important and relevant differences between experimental philosophy as it is traditionally portrayed and the particular application of this approach to folk psychology. In particular, as we will come to see in the coming pages, while the methodology is the same, the purpose or end goal in using the experimental method is quite different when applied to folk psychology. As such, it is important to highlight these differences and distinguish this particular application of experimental philosophy from the broader project. With this in mind, I will first give a history and overview of experimental philosophy in order to illuminate precisely how it has been traditionally deployed and then argue for the relevance of the experimental method and the use of an empirical lens more generally as it relates to the debate regarding folk psychology.

[55] That is, simply using empirical evidence to inform philosophical thinking.

3.2 – The Skeptical and Empirical Foundations of 'Experimental Philosophy'

Experimental philosophers are often critical of what is commonly known as conceptual analysis.[56] As depicted by critics of this practice, conceptual analysis has played a key role in philosophy since at least the time of Plato and is based on the use of hypothetical thought experiments or reflection on our common linguistic or conceptual practices and the intuitions that are generated by these processes.[57] This approach to conducting philosophy is especially popular in epistemology or moral philosophy. In particular, the use of hypothetical thought experiments to reveal our intuitions about the particular application of a philosophical concept is often used in these domains. When we respond to Gettier's (1963) hypothetical thought experiment regarding Jones' car ownership and Brown's geographical position or Foot's (1967) hypothetical scenario of an out of control trolley rocketing down a San Francisco street and we must decide whether to flip the switch to kill the one and save five, we are relying on a hypothetical case and our intuitive response to the case. This intuition is then generally used as evidence for either confirming or rejecting a particular view. For example, Gettier's cases and the intuitions that they produced in readers were used as reason for rejecting the justified true

[56] I use this term as a 'catch-all' for a number of different names or phrases that have been used to identify any approach to conducting philosophy that relies on the use of intuitions regarding concepts, language, and hypothetical thought experiments or any reflective practices on our common or everyday use of concepts, language and thought.

[57] There is, of course, significant debate regarding whether or not conceptual analysis really is the primary methodology of philosophical investigation or whether this practice even occurs. Herman Cappelen (2012), in his *Philosophy without Intuitions*, acknowledges that while most philosophers will agree that philosophy relies on an intuition driven methodology, this just simply is not the case. More specifically, despite the agreement in the discipline, Cappelen argues that analytic philosophers are not guilty of using intuitions as basic sources of evidence. Additionally, Timothy Williamson (2007) in his *The Philosophy of Philosophy*, and elsewhere, has argued extensively that the philosophical method is not different from the methods of other disciplines and that if conceptual analysis is to be rejected then we are left with nothing but a global skepticism about human judgment in general. These criticisms are not insignificant and are worthy of treatment by those experimental philosophers who a) believe that conceptual analysis accurately represents the entire philosophical approach and b) wish to reject this practice outright. My concern here is less ambitious. First, this discussion is meant only to illuminate the history of this approach to conducting philosophy by demonstrating how it is being used and the methods employed. Second, once this elucidation and demonstration is complete I will distinguish the application of this approach in the domain of folk psychology from those just explored. The experimental approach utilized in the later chapters is significantly influenced by the experimental philosophers before me, but I hope to protect myself from some of these global concerns, by demonstrating the particular appropriateness of the experimental methodology in this domain.

belief account of knowledge, and Foot's trolley example could be used to claim that Utilitarianism is the moral theory that best conforms with our intuitive understanding of what is right or wrong. In effect, we sit in our comfortable yet practical armchair and consider a variety of thought experiments to determine whether and when the concept might be appropriately used in order to inform our philosophical theory about the necessary and sufficient conditions under which the concept applies or the rightness of our particular theory. If our thought experiment is constructed in a way that our theoretical commitments say one thing while our intuitions suggest another, then we modify our conceptual understanding or theory in light of this intuition or in some cases we reject our intuition in favour of our definition or theory. Either way, progress is made.

In its earliest form, experimental philosophers were preoccupied with developing a critique of the reliance on intuitions to answer significant normative questions in philosophy. In particular, experimental philosophers used the results of their experiments to support an argument against the view that intuitions are somehow a basic source of evidence that we can appropriately rely on to move the philosophical discussion forward in a positive and legitimate way. More specifically and as we will see in the discussion below, these philosophers wanted to demonstrate that the intuitions produced by these hypothetical thought experiments are more variable, both contextually and culturally, than philosophers tended to assume and that this supported a skeptical argument about the appropriateness of intuitions as a basic source of evidence for a philosophical position. In particular, if we trace experimental philosophy back through the literature, we will notice that it really emerged in response to a skeptical and speculative argument developed by Stephen Stich (1988/1998, 1990) and supported by the social psychological research of Richard Nisbett (see 2003 for a complete collection and review).

Stich was particularly interested in criticizing what he labelled analytic epistemology; effectively any epistemological project that ultimately relied on a conceptual or linguistic analysis to answer important normative matters such as which beliefs or cognitive processes should we have. In his view, epistemologists began with a question about which beliefs we ought to have, and then ultimately ended up relying on intuitions regarding our everyday conceptual practice[58] in order to justify their chosen normative position.

Stich began by noticing that epistemologists answered the question about which beliefs we should have by appealing to a higher order evaluative principle such as 'justification'. More specifically, the beliefs we should have are those that are justified. However, this just pushed the question one level up, as we now need an analysis of which justification practices are the ones we should utilize and which are the ones we should not; in other words, how do we know which justification practices or rules are good ones and which ones are bad? Here Stich takes aim at Alvin Goldman's (1986) appeal to the 'criterion of rightness' to evaluate justification rules as being appropriate to use or not. In particular, Goldman eventually concedes that the question of which 'criterion of rightness' is the one that ought to be used, will be answered by an examination of our "everyday thought or language" (Goldman, 1986, p.58); in effect, a conceptual or linguistic analysis that relies on our intuitions about how a concept ought to be applied. But this, Stich argues, is not a satisfying position. The mere fact that a 'criterion of rightness' accords with our conceptual or linguistic analysis (or intuitional analysis) is not strong enough to answer a very important normative question regarding which beliefs we ought to hold. Stich effectively questions whether we have any reason for thinking that the output of this

[58] Implicit in the reliance on everyday conceptual practices is the assumption that intuitions regarding this practice will be (nearly) universally shared. Historically this has been identified by many philosophers. See Ramsey (1992/1998) for an early articulation of this commitment or Machery, Mallon, Nichols & Stich (2004) for a more contemporary recognition that this commitment underlies this particular methodology.

intuition driven process could have any normative force. His problem with this line of reasoning turned on the issue of cognitive diversity; the idea that cognitive processes could vary for different cultural groups and then, ultimately, that one's intuitions could very well differ depending on one's culture.

The argument Stich provides for the possibility of cognitive diversity can be summarized as follows:[59] Cognitive diversity is a *real* possibility if cognitive processes can, in part, be shaped by one's natural environment. It's worth emphasizing that for Stich, the term cognitive processes would include our "reasoning" abilities, or to put it another way, our cognitive norms. Since cognitive processes are biological processes there is the possibility that they can be shaped by our natural environment, and hence, our culture. Therefore, cognitive diversity is a *real* possibility, that is, not just a logical possibility, but a real world possibility. This is not to deny that some biological processes are not extremely rigid in nature and not open to any variability (e.g. digestion), but there are certainly others (e.g. language) that are open to significant environmental influence and Stich believes that our cognitive processes likely lie closer to language than digestion on this continuum.[60]

While his concern regarding cognitive diversity is speculative, it warrants consideration. If cultural differences lead to cognitive differences it is all too likely that our common everyday

[59] This is a summary and reconstruction of the argument that is developed in Stich (1988/1998). See in particular pp. 95-97.

[60] There is some reason to think that cognitive processing might really be closer to language than digestion. Research in neuroscience, particularly neural constructivism, illustrates the extent to which our brain can be shaped and modified by our environment, showing that our cognitive architecture can be (and in fact is) modified in light of reliable and cognitively essential environmental structures (see for example, Schlagger & O'Leary, 1991; Roe et al., 1990; Quartz, 1999). That is, our cognitive capacities, our knowledge base and more generally, our cognitive architecture are all altered through our constant interaction with the environment. While this research is still fairly new and not endorsed by everyone (e.g. modularity theorists), it would help to support Stich's claim as it situates cognition somewhere between digestion and language, but perhaps closer to language. The point for Stich, however, is simply that it is possible that our cognitive processes differ by culture as this, he believes, significantly undermines the foundation upon which Goldman and others place their confidence in the argument developed.

understanding of some epistemological concept could vary as well. As such, we have not given a satisfactory answer to the question of which beliefs we should hold since different cultural groups could reasonably hold different beliefs that are justified and supported by a criterion of rightness (or some other evaluative measure) that accords with their common linguistic practices, which themselves are the product of cultural influence.

In effect, Stich observes that the normative question simply keeps getting pushed to a higher level until we ultimately rely on our intuitions to justify the story. But analytic epistemologists employing this process have not offered any "reason whatever to think that the notions of evaluation prevailing in our own language and culture are any better than the alternative evaluative notions that might or do prevail in other cultures" (Stich, 1988/1998, p.107) and accordance with the "evaluative notions embraced in our culture" (Stich, 1988/1998, p.107) does not leave us any better off than when we started other than to articulate the structure of our particular linguistic framework. It is perfectly reasonable to think of a scenario where there are two distinct cultural or linguistic groups with coherent belief systems that are ultimately secured or justified by their intuitions about some epistemic concept and relying on intuitions to justify a particular system does not answer the ultimate question regarding which of the two linguistic groups holds the beliefs that we *should* hold.

While Stich's target is simply analytic epistemology, his argument can be applied to any normative philosophical argument that is ultimately reliant on intuitions to make a crucial point in the argument. If intuitions are a product of our cognitive system, but our cognitive system is flexible enough to allow for variation between groups then we can expect that different groups could have different intuitions about core philosophical concepts. In other words, if it is possible to go about cognition differently than we do, then it is possible that our intuitions could be

different than they are and this possibility threatens the legitimacy of the claim that intuitions should be taken as a basic source of evidence for any claim since (1) they could have been otherwise and (2) we are not in a position to choose between competing intuition groups and determine which group should be privileged as right, true or best.[61] We really must come to terms, Stich thinks, with the idea that our cognitive states (broadly construed), processes and structures are really nothing more than "a historical accident" (Stich, 1988/1998, p.106) and so cannot act as evidence to support normative claims about what should be the case.

Stich's argument is certainly a speculative one that relies on the mere possibility of cognitive diversity in order to criticize a particular approach to conducting philosophy. In fact, some critics of his position highlighted this precise fact and wondered whether the logical possibility that Stich identified was possible in this world; whether it is nomologically or psychologically possible (e.g., Pollock and Cruz, 1999). If it is not, that might take some of the wind out of Stich's sail and support the position of Goldman and others like him. In contrast, if cognitive diversity is in fact true in this world, Stich's position could very well be bolstered by this empirical evidence. As alluded to above, there is a growing body of evidence regarding cultural variations in cognition that has been used to support Stich's skeptical argument.[62]

The social psychologist Richard Nisbett (see Nisbett (2003) and Nisbett et al. (2001) for an overview) has collected a great deal of evidence showing that there are cognitive differences between different cultures. Nisbett's studies examine perception, attention, memory, the explanation and prediction of events, the categorization of objects, and even belief revision. Through all of these studies Nisbett has consistently found systematic differences between the

[61] At least not without argument and a defence of one's intuitions that does not end up being circular by appealing to particulars that are granted by the cognitive processes that are under question.
[62] Weinberg, Nichols & Stich (2001) make this exact connection between Stich's speculative and skeptical argument and early research into cultural variation in cognition.

cognitive lives of East Asians and Westerners. The common thread through much of his work is the difference in how Westerners and East Asians attend to or perceive the world of objects and subjects. Most notably, his research suggests that Westerners tend to focus on the agent or the object, while East Asians tend to be more focused on contextual or environmental factors (Nisbett, 2003, p.127). Through a review of all this research and the commonalities he's found through the studies, Nisbett has parsed the general pattern of differences in attention and cognition between Westerners and East Asians into two categories: analytic (for Westerners) and holistic (for East Asians).[63] This body of research effectively confirms Stich's speculation, that the cognitive lives of different cultures can in fact be different, and gives his argument a secure footing to build from.

While one might argue that Stich presented good philosophical reasons for being concerned about the implications of the mere possibility of cognitive diversity for any philosophical argument that is ultimately based on intuitions, the empirical evidence that Nisbett collected really provided the 'meat' to the bones of Stich's argument, suggesting that it is not just a mere fiction. This is precisely the stage at which experimental philosophy was spawned and a significant effort to explore the implications of cross-cultural variation in cognition on intuitions relating to philosophical concepts was undertaken.

3.3 – Experimental Philosophy

Experimental philosophy is a tagline that captures a movement of philosophers who are actively participating in the collection of empirical evidence regarding philosophical matters. More specifically, these philosophers engage in the same process as conceptual analysis by soliciting intuitions regarding hypothetical thought experiments or through reflection on everyday

[63] As many of these studies will be reviewed in detail in Chapter Four, I will refrain from discussing them in detail here.

conceptual or linguistic practices, but experimental philosophers have historically[64] not been interested in the intuitions of philosophers, but rather those of the lay person. Just as philosophers have been doing for millennia, experimental philosophers construct hypothetical thought experiments or vignettes but instead of analyzing them themselves, they present these cases to the lay person to solicit lay intuitions. In order to do this they enter "the field" and actively engage lay people in survey based experiments to collect information from the lay person regarding how they would respond to a particular thought experiment or utilize a particular concept. Importantly, they are not just entering the field willy-nilly, but go in with a particular hypothesis that they set out to test through their experiment. In this way, the methodology of the experimental philosopher is the experimental method.

Often, their motive for engaging in these experiments is to assess whether the standard philosophical intuitions are shared by the folk and more importantly, whether intuitions are subject to variation due to philosophically irrelevant[65] factors such as one's culture. In fact, one of the hallmarks of early experimental philosophy is the ability of experimental philosophers to control for and identify various philosophically irrelevant factors that are associated with particular intuitions. On the basis of these findings, experimental philosophers have typically developed arguments inspired by Stich, attacking the use of intuitions in philosophy.

---

[64] Some experimental philosophers are now turning to explore philosophers' intuitions experimentally. See for example, Sytsma & Machery (2010) as discussed below, who are precisely interested in philosopher's intuitions for the purposes of contrasting them with the folk's intuitions and Schwitzgebel & Cushman (2012) who examined effects of order on the moral judgements of professional philosophers.

[65] I say 'philosophically irrelevant' as experimental philosophers have generally claimed that if conceptual analysis is to lead us to the truth, then intuitions must be a stable and reliable source of evidence that is, for example, not open to manipulation and that leads to repeatable and generalizable results. This is not the case if intuitions vary between cultures, other groups, or for other reasons that should not influence a philosophical claim. Philosophers do not typically think about concepts such as knowledge or reference as being, for example, dependent on one's culture. What constitutes knowledge should transcend these variables and not merely be an artifact of the Western, predominantly male, white and high socio-economic class whose intuitions have led us to define the term in a particular way.

Thus, much of experimental philosophy has been consumed with the critical project of demonstrating that intuitions may not be a particularly good source of evidence when it comes to answering some of the normative questions that plague philosophers and that the early skeptical musings of Stich are in fact supported by empirical evidence. Of course as the variety of philosophically irrelevant factors began to grow beyond just one's culture, the argument against intuitions began to take on a life of its own. That said, the general argument is similar to Stich's original: if intuitions vary depending on philosophically irrelevant factors, then philosophers who rely on intuitions to answer important normative questions will have to give us good reason to think that these intuitions should be relied upon to provide answers to important normative questions in philosophy.[66]

Given the discussion above, it is not surprising to learn that experimental philosophers started by examining cultural diversity in philosophical intuitions. In fact, Stich was a participant in the development of this movement starting in 2001 with Jonathan Weinberg and Shaun Nichols and continues to engage in this sort of research today.[67] These philosophers, turned social psychologists, started with a particular goal in mind, to show that not only is cognitive diversity a reality, as demonstrated by Nisbett's research, but that this diversity is not benign with respect to intuitions regarding philosophical matters. As such, they began by examining the impact of cultural diversity on intuitions regarding philosophical matters.

Stich's original enemy was contemporary analytical epistemology and so Weinberg, Nichols, and Stich (2001) set out to identify whether there were cultural differences in intuitions

[66] Of course, most experimental philosophers would argue that the relativism of intuitions may be rather difficult to avoid. More importantly, that any rationale for privileging one group's intuitions over another will need to be grounded in reasoning that is independent of the particular cognitive processes of that group; otherwise the argument will ultimately be circular.

[67] See for example a fairly recent study on gender differences in epistemological intuitions: Buckwalter & Stich (2010).

to epistemological thought experiments. Not surprisingly, they hypothesized that there would in fact be cultural differences in intuitions regarding epistemological thought experiments and claimed that if this hypothesis is correct then "it is hard to believe that any plausible case can be made for the claim that the normative pronouncements of Intuition Driven Romanticism have real normative force – that they are norms that we (or anyone else) should take seriously."[68] (Weinberg et al., 2001, p.438) Of course, the argument to support this claim relies on the same maneuver that Stich employed in his original formulation, that cognitive diversity challenges the appropriateness of relying on intuitions to answer a normative question since we do not have good reason for privileging one group's intuitions over another. As such, we do not have a satisfactory explanation of why we should accept these epistemologists' answer to the normative question being explored.

Inspired by Keith Lehrer's (1990) work, these philosophers developed a hypothetical thought experiment designed to test intuitions regarding the internalism-externalism debate in epistemology.[69] Weinberg, Nichols, and Stich hypothesized[70] that intuitions regarding the internalism-externalism debate would vary between respondents of different cultures; in particular, those belonging to Western or East Asian cultures. With this in mind, they developed the following *Truetemp* case and presented it to undergraduate students at Rutgers University:

[68] It's worth noting that "Intuition Driven Romanticism" is their (not so friendly) term for any philosophical project that turns on intuitions to make an argument.

[69] Internalism is the view that epistemic justification for knowledge claims should only come from factors internal to the epistemic agent, whereas externalism allows for factors other than those within the agent's scope of introspection as permissible. For example, the reliability of the psychological mechanisms that actually produce the belief, which are external to the agent, are permitted as being relevant factors for epistemic justification on the externalist view, but not the internalist view (see Weinberg et al., 2000, p. 439).

[70] The researchers never actually come out and explicitly state which way they think the results will turn out as they are less interested in the particular way in which intuitions vary and more interested in whether they do vary. That is, it just matters that the results are consistent with the more global hypothesis that epistemic intuitions vary from culture to culture.

> One day Charles is suddenly knocked out by a falling rock, and his brain becomes re–wired so that he is always absolutely right whenever he estimates the temperature where he is. Charles is completely unaware that his brain has been altered in this way. A few weeks later, this brain re–wiring leads him to believe that it is 71 degrees in his room. Apart from his estimation, he has no other reasons to think that it is 71 degrees. In fact, it is at that time 71degrees in his room. (Weinberg et al., 2001, p.439)

Respondents were then asked to state whether Charles "Really Knows" or "Only Believes" that it was 71 degrees in the room. It turns out that one's answer to this question varies with one's culture. While both Western and East Asian respondents were likely to deny knowledge in this case, that is, say that Charles "Only Believes", East Asian respondents were significantly more likely than their Western counterparts to make this assessment.

Similar results were found when the philosophers created a Gettier inspired scenario examining whether the justified true belief account of knowledge is intuitive. With Gettier style cases in mind, Weinberg, Nichols, and Stich developed the following story:

> Bob has a friend, Jill, who has driven a Buick for many years. Bob therefore thinks that Jill drives an American car. He is not aware, however, that her Buick has recently been stolen and he is also not aware that Jill has replaced it with a Pontiac, which is a different kind of American car. (Weinberg et al., 2001, p.443)

Respondents[71] were then asked whether Bob "Really Knows" or "Only Believes" that Jill drives an American car. Again, Western and East Asian respondents show a significant difference in their tendency to assign belief and knowledge to Bob. In particular, while a majority of Western respondents say that Bob "Only Believes", the tendency is reversed among East Asian respondents, where the majority say that Bob "Really Knows".[72] It is worth emphasizing that the

[71] Again, this study appears to have been conducted with undergraduate students at Rutgers University.
[72] This result is consistent with Norenzayan, Nisbett, Smith & Kim's (1999) research demonstrating that East Asians are more likely to make judgements on the basis of similarity when compared to Westerners who are more likely to focus on causation when describing or classifying the world. In this particular case, Weinberg, Nichols & Stich note that "the evidence that *causes* the target to form a belief turns out to be false" which suggests that "EAs might be much less inclined than Ws to withhold the attribution of knowledge in Gettier cases." (Weinberg et al., 2001, p.443)

reversal of the result between cultures is actually fairly dramatic. In the previous *Truetemp* case there was general consensus between the two cultures regarding the attribution of knowledge (that is, while there is a significant difference, generally speaking both cultures withhold knowledge), in this case, however, Western respondents tend to withhold knowledge while East Asians tend to ascribe it. This suggests that these two cultures have significantly different conceptions of knowledge and that one's intuitive judgments regarding Gettier cases is very much dependent on one's culture.

Additionally, Weinberg, Nichols and Stich isolated a population of Indian sub-continent respondents[73] at Rutgers and presented them with the same Gettier case. The effect is even more pronounced in this case, with Indian sub-continent respondents showing an even higher tendency to say that Bob "Really Knows" than either the Western or East Asian populations (Weinberg et al., 2001, p.444). These results are very revealing, as respondents from three different cultures show significantly different tendencies to assign belief and knowledge based on the characteristics of the hypothetical thought experiment. In effect, cultural diversity can sometimes lead to very divergent intuitions regarding epistemic matters.

While Nisbett's research validated the skeptical musings Stich had regarding the diversity of cognitive processes in general, these experiments supported the argument in a more direct way. General cultural differences in cognitive processes present a significant worry regarding the reliance on intuitions in a general sense, but the evidence collected demonstrates clearly that epistemic matters are not immune to this cognitive diversity[74] and encouraged others to look for

[73] Weinberg, Nichols and Stich indicate only that these respondents were of an Indian sub-continent descent. It's unclear whether all respondents are, for example, first-generation immigrants, international students etc.
[74] One could make the argument that Stich's argument was confirmed by Nisbett in the global sense, but that intuitions dealing with epistemological matters would not be subject to this same kind of variation. That is, that there is some reason that epistemic intuitions would be immune to this variation. This idea is similar to the "deep intuitions" idea of Foley (1998).

cultural variation regarding different philosophical concepts or thought experiments. For example, Machery, Mallon, Nichols, and Stich (2006) found similar results when analyzing thought experiments inspired by various theories of reference.[75] As it turns out, Western respondents were more likely than their East Asian counterparts to have intuitions that accorded with a causal-historical theory of reference. It is worth admitting that the results found through this investigation were not as robust as the researchers might have hoped for, indeed the authors admit this.[76] Nonetheless, the conclusion that there is diversity with respect to intuitions is preserved even in light of this slight weakness; the effect need not be extreme to be relevant. In fact, the fact that there is at all a tendency for variation is enough to make the point and hook into the argument that intuitions are subject to cultural influence in unexpected and, arguably, philosophically important ways.

Given these and other investigations[77] it is now generally thought that one's culture can significantly shape many of our intuitions regarding philosophical matters. The precise cases where culture impacts our intuitions about philosophical issues and the degree to which this impact drives a wedge between the different cultural groups is an empirical matter that cannot be answered *a priori*. But the point has been made by these investigations, differences in culture can lead to differences in the intuitions people have regarding philosophical concepts.

[75] The experimental set up was very similar to Weinberg, Nichols and Stich. Undergraduate students at Rutgers and the University of Hong Kong were provided with a series of vignettes and asked probes to assess whether their intuitions about the use of proper names in the vignettes accorded with a descriptivist or causal historical theory of reference. It is worth noting that the University of Hong Kong students were all fully fluent in English.
[76] This could be due to a number of reasons. Perhaps the particular cultural groups selected do not differ in significant ways in terms of how reference operates or perhaps theories of reference are significantly more constrained or deeply held than how lay people attribute knowledge. Alternatively, the particular vignette these researchers developed was quite challenging to understand and that may have impacted the results they found.
[77] We'll see more experiments in Chapter Four pertaining directly to folk psychology.

Research into the variability of intuitions did not end here, however. In fact, Weinberg, Nichols and Sitch hypothesized[78] that intuitions regarding epistemic thought experiments may also vary by socio-economic status (SES). It is a well-known empirical fact that SES is highly correlated with a variety of attitudinal differences and social scientists always track this characteristic as it relates to a variety of social outcomes such as health and often satisfaction with services. As such, it is not a far stretch for philosophers interested in the diversity of intuitions to suspect that SES might also impact intuitions about philosophical matters in some capacity. With this in mind, Weinberg, Nichols and Stich explored the relationship between SES and epistemic intuitions as they relate to Dretske's (1970) *Zebra-in-Zoo* case. The vignette is as follows:[79]

> Pat is at the zoo with his son, and when they come to the zebra cage, Pat points to the animal and says, "that's a zebra." Pat is right — it is a zebra. However, given the distance the spectators are from the cage, Pat would not be able to tell the difference between a real zebra and a mule that is cleverly disguised to look like a zebra. And if the animal had really been a cleverly disguised mule, Pat still would have thought that it was a zebra. (Weinberg et al., 2001, pp.446-7)

Respondents were asked whether Pat "Really Knows" or "Only Believes" that the animal is a zebra. For those who accept diversity in intuitions, it is not surprising that here too there were important variations between different SES groups and attributions of knowledge or belief to Pat. Interestingly, the propensity for a respondent to attribute knowledge to Pat is positively associated with a low SES status (Weinberg et al., 2001, pp.447-8). Specifically, respondents who were categorized as belonging to the low SES category were significantly more likely than high SES respondents to say that Pat really did know the animal was a zebra. This result,

[78] Their hypothesis was supported by some research conducted by Jonathan Haidt regarding differences between SES groups regarding moral intuitions (see for example Haidt, Koller & Dias, 1993).
[79] This vignette was presented to adults who were approached at commercial venues in New Brunswick, New Jersey and who were offered a McDonald's gift certificate for their participation.

experimental philosophers say, is striking since SES should ultimately have no impact on one's philosophical intuitions; from the perspective of traditional conceptual analysis it is not at all clear why one's SES would be a philosophically relevant factor and we should not expect SES to have bearing on whether a philosophical concept like knowledge should or should not be ascribed in a particular situation. As such, the reasoning continues, this observation raises questions regarding the reliance on intuitions to justify a particular philosophical analysis since what constitutes an intuitive response for one SES group may very well be different than another. More specifically, this result means that the intuition methodology would sanction different definitions of knowledge for different groups.

Research into the variability of intuitions then got even more bizarre. Perhaps inspired by the social psychological literature on priming and order effects, experimental philosophers also became interested in the question of whether intuitions regarding philosophical matters would be subject to variation depending on the order in which a series of hypothetical thought experiments were presented. Swain, Alexander and Weinberg (2008) conducted a study[80] to test whether responses to a Lehrer style *Truetemp* case would vary depending on whether the test case came before or after other vignettes. The test case was very similar to the Weinberg, Nichols and Stich *Truetemp* case, but this time respondents were asked to rate their degree of agreement with the statement "Charles knows that it is 71 degrees in his room" on a 5 point Likert scale (Swain et al., 2008, p.154).

In order to test whether presenting a series of hypothetical thought experiments in different orders would have an impact on the *Truetemp* test case, they developed vignettes describing a case of clear knowledge and a case of clear lack of knowledge and asked similar

[80] 220 undergraduate students at Indiana University – Bloomington participated in the study. Respondents were randomly assigned to eight variations of the survey, differing only in the order of the vignettes.

questions using a 5 point scale. Different orders were devised and presented to distinct groups to assess whether there were differences in the willingness to attribute knowledge in the test case depending on the order of cases in the series of vignettes. In fact, this is precisely what they found. More specifically, the experiment found that willingness to attribute knowledge in the *Truetemp* test case was in fact impacted by whether an example of clear lack of knowledge or clear knowledge was presented first. Specifically, willingness to attribute knowledge in the test case was higher when presented with a case of non-knowledge first, but lower when presented with a case of clear knowledge first (Swain et al., 2008, p.144).

This result is consistent with the social psychological literature on priming and order effects,[81] but the work of these philosophers helped to clearly show that intuitions regarding philosophical matters are not immune to this well documented effect. Importantly, these researchers used these results to continue the assault against the use of intuitions in philosophy. In particular, Swain, Alexander and Weinberg note the order in which a vignette is presented to a participant is not a relevant philosophical factor that should impact one's intuitions about the particular case (Swain et al., 2008, p.139), and yet these results demonstrate that this factor significantly shapes the intuitions the lay person reports. This is a striking result and is worthy of emphasis since philosophers often construct hypothetical thought experiments systematically and in a specific order (e.g. "Consider this….what's your intuition? But wait, consider this….."). This observation demonstrates, they note, that intuitions "are susceptible to manipulation" (Swain et al., 2008, p.141) and on this basis they argue that this manipulation significantly undermines the appropriateness of relying on intuitions as a source of evidence in philosophical

[81] The presentation of a case of clear lack of knowledge first, primed the respondents to respond more favourably to the test case. In contrast, presenting the clear case of knowledge first, primed the respondents to be less favourable to the test case.

investigations. In particular, the results of this study led Swain, Alexander and Weinberg to conclude that "[e]vidence so unstable risks being discounted as not truly evidence at all" (Swain et al., 2008, p.141) and so the role of intuitions in philosophy is significantly challenged.

Taken together, these results tell an interesting story of how lay people conceive of, use and attribute philosophical concepts. Unfortunately this story is quite heterogeneous and no clear or comprehensive story appears to emerge. Instead, we find that there are a multitude of variations in intuitions among sub-groups and a variety of factors appear to influence the intuitions people report regarding philosophical concepts. Importantly, while it is true that discovering new and interesting variations in intuitions among sub-groups or identifying factors that impact the precise intuition reported is an interesting and worthwhile pursuit on its own,[82] a proponent of the use of intuitions in philosophy could easily ask "Who cares?" Put more robustly, how does variation or manipulation of the intuitions of the folk at all relate to philosophical attempts to define knowledge or assessments of alternative theories of reference? Why should it matter to the philosopher interested in these matters whether or not there are stable or varied intuitions among the folk?

It is here that experimental philosophers move beyond the merely interesting descriptive project of cataloguing variations and manipulations in the intuitions of the lay person regarding philosophical concepts and argue that taken together these results paint a fairly grim picture for the use of intuitions in philosophy. In order to make this particular move, however, they often hook into or build upon the skeptical musings of Stich to argue that the observed variation or manipulation of intuitions among the lay person significantly undermines the value of

[82] From the perspective of a pure researcher, these results are often quite fascinating and reveal something interesting about the way in which the folk use philosophical concepts.

philosophers' intuitions in answering a normative philosophical question. There are a few possible ways in which this argument can be made.

First, if there is divergence between philosophers' intuitions and folk intuitions then these results demonstrate that philosophers' intuitions could have been other than they are. This possibility may raise doubt as to the appropriateness of relying on a source of evidence that could have been otherwise than it is when trying to answer an important normative question. In particular, Stich's mere speculation that it is logically possible to conceive of a cultural group that forms and revises beliefs in significantly different but internally consistent[83] ways, while not explicitly verified by these findings, is supported by them.

Second, they argue that these results demonstrate that there does not appear to be a single common sense or everyday notion of philosophical concepts that can be relied upon given the significant variability these experiments highlight. This, they claim, is particularly problematic for any philosopher who follows Goldman and appeals to a notion of everyday thought or language to justify their particular philosophical position since they will have to contend with different and competing groups. Additionally, if the philosopher expects that their intuitions will be universally shared, these results cast doubt on this assumption as well.[84]

Third, and perhaps more importantly, if philosophers intend to reject common sense intuitions as "untutored" and argue that their "expert" intuitions can be preserved, they will have to contend with the fact that these "expert" intuitions may be nothing more than a product of their own particular group which, it is worth nothing, is predominantly male, white, Western in culture, and of a high socio-economic status. As such, philosophers will need to provide good

[83] Internally consistent insofar as the justification practices or criterion of rightness that is used to assess the belief formation or revision process accords with their own everyday thought or language.
[84] As noted above, the appeal to everyday thought or language may include the assumption (implicit or explicit) that there is a single universally shared notion that is being appealed to.

reason for privileging their intuitions over others and for rejecting the idea that intuitions should accord with *everyday* thought or language. However, in crafting this reply philosophers will need to be particularly cognizant of the fact that their group is one that is highly trained and intensely selected for and that these intuitions may be less reflective but rather more reinforced by our community (Machery et al., 2004, p.B9). As such, there may be good reason to worry about the reliability of this group's intuitions.

In effect, this negative and critical project has used experimental evidence to challenge philosophers who wish to rely on intuitions in their practice to provide evidence or good reason to think that the intuitions of philosophers should be privileged over those of the folk, that they are not subject to the same types of variations, and that they can be relied upon safely in our pursuit of the truth. Not surprisingly, many philosophers have responded and taken up this challenge. In fact, objections to the relevance of potential or real variation in what constitutes everyday language or thought have been considered for some time.

Foley (1998) has suggested that there may be "deep intuitions" that are not subject to the challenge of cognitive diversity that Stich raised. The idea here is that some intuitions are not just confidently held, but are deeply held such that critical reflection will not change or alter them, that is, they are "reflectively stable" (Foley, 1998, p.244). Presumably, if anyone is to have these deeply held intuitions, it might be philosophers, but all the better if there is cross-cultural agreement among folk intuitions. Bealer (1998) argues that there is a modal or logical tie between our intuitions under sufficiently high cognitive standards and the truth. This admits that intuitions in general are fallible, but that under the right cognitive circumstances they can continue to play a role in philosophy without risking failure at every turn. Presumably, the

suggestion is that philosophers are likely or more likely than the lay person to be able to reach this sufficiently high cognitive standard and access the truth using intuitions.

Others have taken a more direct route to privileging philosophers' intuitions, carving them off from the instability and variability found in folk intuitions. Hales (2003) argues that philosophers' intuitions are more reliable than folk intuitions because years of training have provided us with access to higher cognitive circumstances. In effect, like physicists who have better intuitions about physics, philosophers have better intuitions about philosophical problems (Hales, 2003, p.5). Similarly, Ludwig (2007) argues not that philosophers' conceptual competence is better than the folk, but that there is a difference in performance that is being identified in these experiments. Put another way, there are performance errors in the intuitions of non-philosophers whereas philosophers considering the same thought experiment are better at reporting intuitions that align with the underlying conceptual competency.

While I think experimental philosophers and skeptics alike are well positioned to construct replies to these objections,[85] there is some merit to these concerns, especially as it applies to the domains of philosophy explored thus far. It might be possible to argue that while the term "knowledge" applies broadly and beyond just the philosophical domain, that there is room to create a divide between the common sense and philosophical conceptions of knowledge. The technical definitions we philosophers develop and provide could possibly be something separate from or at least distanced from common sense conceptions[86] of the same concept. For instance, we're not interested in when the folk attribute this concept, but rather when the concept

[85] Some of these objections are addressed in the research we've reviewed thus far. See for example, Machery, Mallon, Nichols and Stich (2006); Stich (1988/1998, 1990); Swain, Alexander and Weinberg (2008); Weinberg, Nichols and Stich (2001). This is not to say, however, that there is not a lively debate over these objections and the replies that are offered.

[86] I am cognizant that some philosophers (e.g. Goldman, 1986) appeal to *everyday* thought or language to justify their position. I'm simply acknowledging here that this may not need to be the case and in response to the empirical challenges to this maneuver, it may be appropriate to appeal to a more restrictive class of intuitions.

should be attributed. There is a sense in which the concept of knowledge that philosophers are really interested in just is a philosophical concept. Alternatively, there is also some technical knowledge or skill that is required to properly assess challenging philosophical issues and this may have bearing when assessing the relevance of untrained folk intuitions. For example, bioethical issues are often extraordinarily difficult and it takes a moral philosopher (many would argue) to effectively navigate these challenges because our understanding of morality is at least somewhat departed from common sense conceptions of morality in important and relevant ways. For example, we are better able to recognize, apply, and balance various principles (e.g. beneficence and autonomy) in order to assess whether a particular action is morally permissible or not.[87] In this sense, the skill and training that comes with philosophical training may be directly pertinent when trying to answer important normative questions, even when relying on intuitions to help answer this question.

I raise these issues not to address them at length,[88] but to demonstrate that there is a significant and important debate as to the conclusions that experimental philosophers draw on the basis of the experimental evidence. That is, while the descriptive project of cataloguing folk intuitions is valuable and important, there is still significant debate regarding what to do with these results. Rather, the purpose of this section and the chapter more broadly has been to explore experimental philosophy to discover the general methodology experimental philosophers employ and to reveal the purpose under which much of this work has been undertaken.

[87] As a practical example, hospitals often have a professional bioethicist on staff to help physicians, patients, and families navigate complex ethical issues that arise in end of life decision-making. Clearly these professionals possess a skill set that positions them in such a way that is different from physicians, patients, and families and that is seen as helpful.

[88] A full and appropriate treatment of this debate is well beyond the scope of this project and is not crucial to the argument of this chapter.

The discussion above demonstrates a clear adoption of the experimental method to develop and design tests to critically evaluate philosophical claims. Experimental philosophers are not interested in merely theorizing about matters, but actively engaging in the production of empirical evidence that is designed to help us build a philosophical argument. No more are philosophers passively waiting for empirical findings from other disciplines, but we are actively taking up these efforts and designing the studies to directly suit our purposes. Importantly, early worries about methodology[89] (e.g. sample size, survey design, data analysis, etc.) are being put to rest by continually improving the experimental set up and working with psychologists, statisticians and other researchers to develop robust studies. But this discussion also clearly illuminates the primary goal of many of the early experimental philosophers. In particular, they were producing this empirical evidence to build and support an argument against the use of intuitions in philosophy. In particular, they have used the results of their experiments to support a challenge against this practice by identifying variability and instability in intuitions and used this to challenge some of the assumptions of those philosophers relying on intuitions to answer important normative questions.

Because much of the work that follows in Chapter Four and Five falls under the same experimental philosophy banner I have taken the time to explore this movement in detail. However, I think there are important and significant differences in the applicability and purpose of experimental philosophy when applied to folk psychology when compared to, for example, epistemology. Put simply, most of the work just explored can be portrayed as the negative use of experimental philosophy, whereas there is also a positive application for experimental philosophy. The commonality lies in the methodology, but there is divergence in terms of the

[89] See for example Weinberg, Nichols and Stich (2001) for recognition that some critics expressed a worry about the methodology.

relevance and applicability of experimental philosophy to each debate. In particular, while those experimental philosophers just explored have to go to some length to defend the relevance of their findings and philosophers may be able to drive a wedge between philosophical and common sense notions of a concept in response to this evidence, the relationship between folk psychology and experimental philosophy is much tighter. First, like much of philosophy of mind, there is an immediate and essential connection between the philosophical and common sense conceptions of folk psychology. That is, the object of philosophical interest *just is* the common sense notion of folk psychology.[90] Second, much of the work regarding folk psychology is effectively a descriptive project that aims to understand exactly how folk psychology is used by the folk. That is, philosophers are not interested in some important normative question about how one ought to navigate the social domain, but rather how exactly it is that we do navigate the social domain. Importantly, experimental philosophers interested in folk psychology have the exact same purpose; they are also interested in the descriptive project of understanding how people think about people. In this way, and as we will see in more detail below, the relevance of experimental philosophy to folk psychology, and really philosophy of mind more broadly, is markedly different than the project that we have just explored above.

3.4 – Folk Psychology and the Experimental Method

The philosophy of mind is similar to other areas of philosophy in that it typically relies heavily on thought experiments. No discussion of the possibility of artificial intelligence would be complete without considering Searle's (1980) *Chinese Room* thought experiment as a probe designed to illicit intuitions denying the plausibility of a rule following system possessing

[90] This may also be true for moral philosophy as well. For example, moral philosophers may tend to prefer a normative ethical theory that is closely related to common sense morality and so in this sense, these philosophers may too be very much interested in folk morality.

understanding. Similarly, no discussion of functionalism would be complete without considering Block's (1978) *China Brain* thought experiment as a way to highlight the counterintuitive and potentially implausible implications of multiple realizability. In fact, I suspect you would be hard pressed to pick up an introductory textbook to philosophy of mind that was not riddled with thought experiments designed to probe for particular intuitions as evidence for or against a broader theory of mind, consciousness, perception etc.

However, philosophy of mind may be different from many other areas of philosophy in that it is often more closely associated with common sense issues and our own personal experience with being a mental, conscious thing. If our intuition probes are about "what it is like" to have a mental experience, this is something we all share and we all directly experience. So when Chalmers (1995) writes about the qualitative experience associated with the whir of information processing his cognitive system is undergoing, his claim that "[t]his subjective aspect *is* experience" (Chalmers, p.10, emphasis added), is a claim not about his own experience as a trained philosopher, but the experience of all of us where we all have access to our own subjective experiences. Similarly, when Searle (1994) asserts that one of the "simple and obvious truths about the mind" is that "[w]e all have inner subjective qualitative states of consciousness" (Searle, p. xi) he is, again, not making a statement only about his own inner subjective qualitative experience, but is making a claim that "we all" have *this* experience. Presumably then, when a hypothetical thought experiment is constructed to solicit intuitions regarding conscious experience, perception etc. it is meant to solicit the same intuitions from all of us who share in this conscious experience.[91]

[91] Of course it could be argued that there is some skill involved in analyzing one's own subjective experience carefully enough to understand and notice the fine distinctions philosophers tend to make. Indeed there is some evidence that introspective training leads to improved accuracy with respect to some elements of subjective experience (see for example, Fox, Zakarauskas, Dixon, Ellamil, Thompson, & Christoff, 2012). However, it is not

Justin Sytsma and Edouard Machery (2010) have collected a catalogue of quotations, including those above, regarding subjective experience to demonstrate that there is an obvious linkage between the philosophical and the common sense conceptions of mental experience. That is, that philosophers explicitly and purposely make reference to common sense conceptions of mental experience in their philosophical endeavours, stating that their project is a common sense project. For example, Block (2004) refers not to a philosophical conception of subjective experience, but rather the "common-sense conception of subjective experience" (p.785). Similarly, Dennett (2005) speaks of our folk theory of consciousness as something that is picked up "in the course of our enculturation" (pp.26-27), and Patricia Churchland (1988) explicitly states that she is interested in "the old folk notion of consciousness" (p.302), while even Alvin Goldman (1993) sets his sights on "the folk-psychological notion of phenomenal consciousness" (p.364). These quotations are striking in that they immediately and inseparably tie the philosophical and common sense projects together. The object of interest for these philosophers is not just some philosopher's creation, but a concept, theory, or experience we all share and all have.[92]

clear whether the results of the aforementioned experiment apply to the entire domain of subjective experience or just the one element of experience examined in the experiment, namely introspective awareness of tactile sensitivity in the body. Additionally, if we admit that skill is in fact associated with improved accuracy or understanding with respect to one's subjective experience this could mean that we are now no longer trying to describe a common sense experience. Indeed, training or skill to this degree is exactly what the 'common sense' or 'everyday' designation is supposed to avoid. Whether the wedge between a skilled and common sense notion of experience has philosophical relevance is a question beyond the scope of this project and is not crucial to the more central claim of this section; that when trying to understand folk psychology we cannot try to carve our project off from the common sense project. I will, however, concede that it could turn out that subjective experience is less strongly associated with the common sense project than folk psychology is.

[92] It might be worth noting that many of these philosophers rely on their intuitions or their personal subjective experience to answer some important normative questions about subjective experience or consciousness in general, but that they are doing this in a way that necessarily ties their project to the common sense notion of subjective experience. In this way, the intuition or reflective practice plays the same role as Goldman's appeal to everyday thought and language. What is important is that unlike epistemologists who might ultimately drive a wedge between the common sense and philosophical notions of knowledge, these philosophers of mind are committed to the common sense notion in a way that makes driving a similar wedge less plausible. This will become more evident in the discussion that follows.

Sytsma and Machery catalogued these quotations to demonstrate that with respect to subjective experience or consciousness, philosophers of mind have effectively stated that what they are talking about is the same thing the folk are taking about and that there is an expected overlap in terms of what they are talking about and what the folk would say about the same concepts or experiences. On the basis of this expectation, Sytsma and Machery argue that there is value in assessing empirically whether the philosophical and folk concepts of subjective experience are aligned (Sytsma and Machery, 2010, pp.300, 320-1). The purpose here is not to question the use of intuitions in general,[93] but rather to question whether or not the intuitions of philosophers line up with the intuitions of the folk. Of course, the expectation is that if philosophers are speaking on behalf of everyone, accessing a common sense conception of subjective experience, then there will be alignment between the philosophers and the folk since we are talking about the same thing and should generally have the same kinds of experiences.

In case the quotes and discussion above are not sufficient to demonstrate that philosophers are really making a claim about the folk notion of subjective experience, Sytsma and Machery invited philosophers[94] to participate in one of their experiments to test this claim. Upon completing a series of questions regarding subjective experience, philosophers were then asked whether they thought the folk would respond in the same way. Not surprisingly, given the above quotations, the philosophers in this study generally did think that the folk would give answers to the questions that accorded with their own answers (Sytsma and Machery, 2010, p.308). This result confirms the general expectation that when philosophers of mind are talking about subjective experience they are often explicitly and purposely talking about a common

[93] Although Machery has been known to do so (see Machery, Mallon, Nichols & Stich, 2004).
[94] Defined as those who were completing or had completed an undergraduate degree with a major in philosophy or at least some graduate level courses. They also note that when the results were analyzed for only those with graduate training or higher, the results were unchanged.

sense conception and that they expect their responses to truly align with the folk conception. As such, philosophers think of themselves as employing their own intuitions to discover the common sense notion of subjective experience.[95]

The situation is very much the same and in fact much simpler when examining the discussion surrounding folk psychology. Because of the worries typically associated with experimental philosophy, Sytsma and Machery wanted to collect quotes to clearly demonstrate that the notion of subjective experience that was being analyzed by philosophers was the folk notion in order to prevent critiques from driving the type of wedge I identified above in my discussion of knowledge ascriptions. This connection between the philosophical and folk project is much more obvious when examining folk psychology than subjective experience, in part, because the term 'folk psychology' itself makes the connection that Sytma and Machery had to discover entirely clear in this case. As was made clear in Chapter Two, philosophers really do conceive of folk psychology as a 'common sense psychology' that is used by not just philosophers, but really by humanity more broadly. Fodor's claim that "common sense psychology works so well it disappears" (Fodor, 1987, p. 3) is not a claim about philosophers' ability to predict and explain the behaviour of other philosophers, but a claim about a *common sense* understanding of others that is used by the folk to predict and the explain the behaviour of themselves and others. Similarly, while he aims to argue that the practice is fraught with problems, Churchland sets his target squarely on a "commonsense conception of psychological phenomena" (Churchland, 1981, p.67) anchoring his project to the "average person" (Churchland, 1981, p.68). Later he refers to folk psychology as being "a framework of concepts,

[95] For the interested reader, it is worth noting that the intuitions do not align. The results of Sytsma & Machery's (2010) research suggests that the folk and philosophers have significantly different conceptions of subjective experience.

roughly adequate to the demands of *everyday life*" (Churchland, 1989, p. 225, emphasis added). Analogously, Dennett (1987) refers to folk psychology as something "familiar to us since childhood and used effortlessly by us all every day" (p.7). Again, he is talking about all of us, not just philosophers. As a final example, in Goldman's (2006) defense of a hybrid view of simulation theory he sets his sights on everyone, noting that "[p]eople attribute to self and others a host of mental states" (Goldman, p.1). While examples could be produced almost *ad nauseam* two more taken from the introductory texts on the philosophy of mind that we explored in Chapter Two will make the point:

> "Much of our ordinary psychological thinking and theorizing ("commonsense" or "folk" psychology) involves propositional attitudes; we make use of them all the time to explain and predict what people will do." (Kim, 2006, p.15)

> "'Folk psychology is a conceptual framework' and/or 'network of principles' (perhaps largely implicit) used by ordinary people to understand, explain, predict their own and other's behaviour and mental states" (von Eckardt, 1994, p.300)

The language could not be more plain. Summarizing decades of research on the topic and presenting this view for a new breed of philosopher, Kim unequivocally refers to folk psychology as a common sense theory and von Eckardt explicitly refers to 'folk psychology' as being used by "ordinary people".

These quotations are analogous to those catalogued by Sytsma and Machery regarding subjective experience. Again, they are striking in that they immediately and inseparably tie the philosophical and common sense projects together. As such, the object of interest is not some philosopher's creation, but a common sense theory that we all share and all use. These philosophers, who are fairly representative of the prominent and traditional construal of folk psychology, have all made an expressed commitment to trying to understand the common sense

conception of folk psychology and thus, a commitment to articulating the nature of folk psychology as practiced and used by the folk.

This commitment is also demonstrated when we observe that much of the debate around folk psychology has been a debate regarding the cognitive mechanisms that underlie the folk psychological framework. That is, much of the debate regarding folk psychology has been over the attempts made to articulate the psychological mechanisms by which we successfully utilize our folk psychology. Traditionally, as we saw in Chapter Two, this has meant postulating the mechanisms by which we predict and explain the behaviour of others using folk psychology. Whether the mechanisms posited are founded in a theoretical analysis or based on a simulative account of folk psychology, there is a clear commitment to the attempt to understand exactly how it is that the folk engage in the practice; it is a claim about the actual practice of the folk. Therefore, these philosophers are not merely speaking about a philosophical concept, framework, or theory that the folk *may* be using or that they *ought* to be using;[96] they are not merely interested in some philosopher's term of art. They are interested in and speaking explicitly about the *folk* concept, framework, or theory.

Similar to the analysis of Sytsma and Machery, with respect to folk psychology the philosopher's reflection on this practice[97] or their intuitions about some folk psychological concept, is meant to say something about the practice itself as used by the folk. As such, there is an expected overlap in the judgments philosophers make regarding folk psychology and those

[96] Although understanding what framework *ought* to be used may be a worthy project, it just simply is not the project these philosophers are interested in.

[97] I should note that it's not really clear whether the word 'intuition' appropriately applies in the domain of folk psychology. It might be fairer to say that philosophers are merely reflecting on their own practice with folk psychology and reflecting on their observations of others. As such, it may not be appropriate to use the word intuition. Because I have used this term above and because this term has been used throughout experimental philosophy and philosophy more generally to mean a variety of things, I will continue to use it with the caveat that with respect to folk psychology it may not be the most suitable word.

that the folk would make in similar situations. I admit, however, that I do not have empirical evidence like Sytsma and Machery that explicitly catches philosophers expressing this expectation.[98] However, the quotes above and the debates regarding folk psychology demonstrate this commitment clearly enough to continue. More importantly, if this claim is true of subjective experience I cannot imagine how it would not be true of folk psychology since the former is a much more challenging and difficult notion to understand and the latter really is a common, everyday practice. To summarize, philosophers interested in folk psychology are making claims that are meant to illuminate something about this real and practically utilized framework. In this way, the philosopher interested in folk psychology is very much engaged in a descriptive project about how the folk do navigate the social domain.

This commitment aligns the philosophical project with the common sense project in a way that gives credence to the use of experimental philosophy to investigate whether or not these two projects are in fact aligned; whether the expected overlap in the descriptive project is in fact present. Experimental philosophy applied to the domain of folk psychology, as we will see in the coming chapters, is generally aimed at understanding how the folk (not philosophers) use folk psychology. In particular, experimental philosophers frequently develop studies to explore the attributive practices of the folk when it comes to core folk psychological concepts such as belief, desire, causal responsibility, and intention. Again, the goal here is not to question the use of intuitions in general, but rather to explore how the folk utilize this framework and to then draw conclusions regarding the nature of the framework on the basis of this investigation. In effect,

[98] I do have anecdotal evidence in the opposite direction, however. Upon completing one of the surveys that I will review in Chapter Five, one of the respondents approached me to ask more about the research. Upon explaining the philosophical point of view and how I was testing this view through the survey he answered, the respondent was honestly surprised that any philosopher would say anything different than his response to the survey. In other words, this respondent expected philosophers to embrace the folk understanding of the concept in question.

experimental philosophy applied to folk psychology really is very much just like a social psychological endeavour aimed at understanding how the folk use folk psychology and understand one another. Because experimental philosophy helps to reveal the nature of folk psychology, or at least some aspects of it, we can use these results to explore whether there is in fact the expected overlap between the philosophical and common sense conceptions of this common sense practice. In other words, because the philosopher interested in folk psychology is engaged in a descriptive project, and because experimental philosophers interested in folk psychology are also engaged in the same descriptive project, there is significant value in exploring how experimental philosophy can be used to investigate folk psychology. In effect, experimental philosophy helps us to empirically verify (or discredit) the armchair musings of philosophers interested in this real world practice. Without evaluating the philosophical project through this lens, how are we to be at all confident that we have actually come to understand how the folk understand each other?[99] After all, it is not guaranteed that what the philosopher says and what the folk do will be aligned,[100] or at least this is an empirical question that needs to be answered by empirical means.

[99] One commentator has noted that experiment is not necessary so long as you know how to use the word. So when psychologists investigate memory, they start with the folk concept without first soliciting intuitions about this concept. I think there is some merit to this concern, but I think that the psychologist's position is somewhat different than the philosopher's. While psychologists may start with a folk concept without soliciting intuitions about the concept itself, they are often engaged in an empirical endeavour that is very much anchored to the folk concept of memory (insofar as the folk are participants in their research). As such, I suspect that if the psychologist's version of the folk concept was significantly out of line with the way in which the folk were using the term, they might come to discover this through their empirical endeavours. Moreover, I suspect that psychologists would be open to modifying their concept of memory if it significantly departed from the folk concept, or at least recognize that they are now talking about different concepts. The situation is very different for philosophers interested in folk psychology. They are not (typically) engaged in empirical investigations regarding how the folk use the framework and it is not clear whether they *can* separate their project from the folk project or whether they would be willing to revise their theory on the basis of empirical evidence challenging their view.

[100] If anything, all of the experimental evidence collected and recounted in the earlier sections of this Chapter detail that often philosophers and the folk do not agree or at least there are significant variances between them. What a philosopher finds to be intuitive may not actually be intuitive to the folk.

It should be clear then that there are important differences between experimental philosophy applied to folk psychology and experimental philosophy as traditionally portrayed and as explored above. The purpose of exploring the results of experimental philosophy as it pertains to folk psychology is not to question the general use of philosophers' intuitions in answering important normative questions such as how we *should* go about attributing beliefs and desires. Moreover, the purpose of exploring the results of experimental philosophy as it pertains to folk psychology is not to question the use itself of philosophers' intuitions regarding folk psychology. This project is not an inherently skeptical project with the aim to significantly revise the philosophical method. Rather, the purpose of exploring the results of experimental philosophy as they pertain to folk psychology is simply to test whether the intuitions of the philosopher match those of the folk. Or put another way, to see whether the descriptive project that these philosophers have undertaken is at all reflective of the practice it is supposed to be capturing and if not, ultimately help to formulate an accurate view of the practice. As it applies to folk psychology, experimental philosophy is really much more of a positive project.

Above I questioned the relevance of variation and instability in folk intuitions when trying to answer tough normative questions like which beliefs we should have. I noted that it may be possible to drive a wedge between, for example, the philosophical and folk concept of knowledge. In particular, when thinking about what ought to be the case I noted that philosophical training and skill might actually be relevant considerations. However, a similar wedge cannot be driven between the philosophical and folk notion of folk psychology. To put it as plainly as possible, the whole discipline has the term *folk* squarely positioned in the name of the framework. If philosophers do seek to distance themselves from the common sense project, then they have effectively removed the *folk* from folk psychology and would ultimately be

investigating some version of a *philosopher's* psychology.[101] This may be a fine move in principle, but it would significantly change the discourse and reframe the object of interest to something much different than what philosophers have been claiming to study. Moreover, taking this move seriously would significantly undermine much of the debate regarding folk psychology that has been undertaken for the past few decades. As such, this move might ultimately be a foolish one as it pushes us into a new territory that bears nearly no relation to where we started and why we engaged in this discussion in the first place. In this respect, our investigations into folk psychology are intimately and necessarily tied to the common sense concept, framework or theory. Importantly, this difference between folk psychology and other domains of philosophy firmly secures the relevance and applicability of experimental philosophy to the discussion. Philosophers interested in folk psychology are effectively engaged in an attempt to reveal the nature and structure of a real-world framework or, to borrow a phrase and analogy, a "proto-scientific project modeled on the Chomskyan tradition in linguistics" (Machery et al., 2004, p. B9). The goal is not to prescribe a particular framework, but to discover the framework and experimental philosophy helps us to assess whether the model created is empirically adequate.

3.5 – The Empirical Lens

The use of experimental philosophy to understand folk psychology is really just one component of the broader empirical lens through which philosophical analyses of folk psychology can and should be evaluated. Thus, the results of experimental philosophy are no different than any other source of empirical evidence that is used to support, reject, or test a particular conception or view of folk psychology; it is no different than, for example, the relevance of psychology, and in

[101] This is analogous to the skilled or trained version of subjective experience mentioned above.

particular cognitive psychology and social psychology, social science, and neuroscience to the discussion of folk psychology.

The use of an empirical lens in general, however, comes with a whole lot less baggage than experimental philosophy. That is, the particular type of experimentalism, the radical nature of experimental philosophy, the 'new car smell', and the typical ways in which it is employed make it worth a significant treatment that the general use of an empirical lens does not need. In contrast, the use of empirical evidence in philosophy more generally is quite frequent and is not lost on all philosophers. This is especially true for those working in the philosophy of mind and then again, especially among those working on issues related to folk psychology.

In fact, using well established domains[102] of empirical inquiry to investigate, support, or reject particular views about folk psychology is quite common practice among philosophers. Fodor (1998) argues that a computational account of cognition, or the representational theory of mind, will vindicate his account of folk psychology; firmly putting his faith in cognitive science and psychology and relying on the empirical sciences to support his view. Whereas in his rejection of folk psychology as a viable theory about the mind, Churchland relied first on neuroscience (1981) and later on the connectionist paradigm that emerged in cognitive science and artificial intelligence research (1989). More directly, proponents of simulation theory relied extensively on developments in neuroscience to argue in favour of their view. In particular, Gallese and Goldman (1998) attempted to demonstrate how mirror neurons[103] provide scientific

[102] As noted above, this could include psychology, and in particular cognitive psychology and social psychology, social science, and neuroscience and is in contrast to experimental philosophy which is still fairly new and still developing in sophistication and approach.

[103] Mirror neurons are those neurons in the brain that are activated both when witnessing someone else's behaviour and when engaging in that behaviour myself. A great example is when pain neurons fire in our brain while we watch someone else get injured. What is perhaps most interesting about these neurons is the fact that while they fire, the same behaviour is not produced in us. So while I see someone get hit in the arm, neurons in my brain will fire as though I've been hit in the arm there is no pain reaction associated with the firing of these neurons.

vindication of the cognitive mechanisms required for simulation theory; in particular, empathy. But proponents of a theory-based analysis of our folk psychological competence also appealed to research in developmental psychology, including that of Wimmer and Perner (1983) or Alison Gopnik and colleagues (e.g. Gopnik and Wellman, 1992), to demonstrate that children operated much like a little scientist navigating the social world by constructing theories about mental states. Moreover, philosophers often engaged with the empirical research of Premack and Woodruff (1978) and their attempt to answer whether chimpanzees have a theory of mind. More recently, Nichols and Stich (2003) relied extensively on Alan Leslie's (see for example, Leslie, 1987) research into pretence to help them confirm that the hybrid (simulation and theory based) cognitive architecture they postulate is correct. But Goldman (2006) articulates the expectation that philosophers interested in folk psychology must appeal to empirical evidence most strongly stating that "[s]erious students of the field cannot ignore any of these disciplines, because each contributes valuable insights" (p.4) and includes philosophical, psychological, and neuroscientific evidence into his account of folk psychology.

Each of these examples clearly demonstrates that philosophers interested in folk psychology are not afraid of and often welcome the opportunity to appeal to empirical evidence in their investigation. This is likely due to the descriptive nature of the project where there is significant value in ensuring that their armchair reflections and reasoning align with empirical evidence that is being collected regarding the practice of describing and interpreting the behaviour of others. As such, the use of the empirical lens in general is not unusual when evaluating the claims of philosophers regarding folk psychology.

In this way, my use of the empirical lens in the following chapters is not new and should be fairly non-controversial since it has been used by many before me. Most importantly, as I

argued above it is an empirical question whether the philosophical and common sense projects are aligned and the empirical lens will help us to answer this question. This applies whether the source of evidence is experimental philosophy or the more established domains of empirical inquiry I spoke of above. What is new, however, is the question that I'm trying to answer by using the empirical lens. While the philosophers quoted above have relied on a variety of sources of empirical evidence to evaluate the mechanisms that underlie our folk psychological competence, I am using the empirical lens to investigate a question that is one step prior in the investigation. That is, I shift the focus one step back and use empirical evidence to test the assumptions under which these philosophers have been working: assumptions regarding the purpose and goals of the folk psychological practice. It is this shift that is fairly new and where some resistance may be felt. But just as it is an empirical question whether our folk psychological practices are supported by a theoretical competence or simulation based mechanism, whether the goal of folk psychology is to provide quasi-scientific explanations and predictions of behaviour is also an empirical question. In other words, it is an empirical question whether or not the folk use folk psychology with the same purpose that philosophers say they do.[104]

As such, the use of an empirical lens to evaluate the assumptions that underlie the traditional and prominent construal of folk psychology is consistent with the approach taken by other philosophers interested in folk psychology. What is new in what follows is the question that is being analyzed and the use of experimental philosophy to supplement the investigation. However, in both cases there is nothing inappropriate about the move. Experimental philosophy

[104] It is worth emphasizing that this is not to say that we should merely ask the folk what the purpose of folk psychology is. But rather, to examine empirical evidence regarding the practice itself and assess whether this empirical evidence is consistent with the presumed goals of folk psychology. Put another way, we should empirically investigate the practice itself and see what this reveals about the function of folk psychology.

is just a subset of the broader empirical lens, and using empirical evidence to evaluate claims regarding folk psychology is entirely appropriate and common.

3.6 – Conclusion

I've said all of this with great force, as I am concerned that there will be some resistance among more traditional philosophers regarding the extent to which I am relying on empirical evidence and the types of empirical evidence that I am relying on. Much of what I have argued for will come as no surprise to some readers, in fact, it may be completely obvious. If that's the case, all the better for what follows. However, given that the goal of this project is to significantly undermine core assumptions that are held and have been held by a significant number of philosophers for some time, it was necessary to take the time to justify the use of an empirical lens in this particular philosophical investigation. Moreover, given that experimental philosophy will play a central role in this evaluation and it is often viewed negatively by many philosophers, it was necessary to take the time to both examine this methodology in general and then justify the applicability of this approach to the domain of folk psychology.

In closing, it is worth noting that while experimental philosophy is often associated with a very specific skeptical and negative project and that while my project is primarily a critical one, the empirical evidence explored in the coming chapters can be used to help shine a light on the positive project that stems from this work. In other words, the empirical evidence we will explore in the coming chapters provides the foundation for a positive story that has yet to be fully developed in the literature. It forces us to conceive of folk psychology in a new light, and provides us the foundational evidence we need when developing this new theory.[105] At this point

[105] This again separates experimental philosophy as applied to folk psychology from the experimental philosophy project we explored above. Most notably, some critiques think that if the arguments of experimental philosophy are correct that this risks a form of global skepticism about human knowledge (see for example Williamson, 2007). In

we've reviewed what philosophers think about folk psychology and why empirical evidence is so important to this project. So, without further ado, let's look through the empirical lens and find out how the folk use folk psychology.

response to this concern, some experimental philosophers have made an attempt to limit the implications of their project and avoid this extreme of a conclusion (see for example Weinberg, 2007).

Chapter Four: The Psychological, Social & Cultural Underpinnings of Folk Psychology

4.1 – Introduction

In the previous chapter I set the stage for investigating folk psychology empirically and in particular, investigations done under the umbrella of experimental philosophy. These studies will be explored at length in the next chapter as they typically focus on the operation of individual folk psychological concepts. In this chapter I want to start with a wider look at the empirical evidence available to philosophers interested in folk psychology. In particular, the evidence we will explore in this chapter comes from a range of disciplines including social psychology and even behavioural economics. The aim of this chapter is to highlight two often related findings that emerge from an empirical investigation of our folk psychological and related practices. First, the results explored below suggest that we are not nearly as good as we (and philosophers) think we are at predicting and explaining behaviour. Second, and more importantly for the purposes of this project, it appears that there are a variety of considerations beyond the scope of quasi-scientific explanation and prediction that influence the way we construct, use and digest folk psychological stories.

We begin with some historical examples from psychology and behavioural economics that highlight psychological biases that impact the way we recall and utilize information and that affect our decision-making. I present these examples as fuel for the fire that is going to grow over this and the next chapter, to introduce the reader to the ways in which our folk psychological practices might fail or be open to influence, and to set the stage for a review of the confabulation literature. Here we'll discover that our folk psychological competence is not always equipped to identify the underlying causes of our behaviour. However, in spite of this failure, we confidently offer up robust and creative explanations of our behaviour. Importantly, the fact that we offer these fictions as explanations creates a gap in our understanding of what

exactly people are doing when they explain behaviour. What leads us to fail to provide accurate explanations and yet provide creative and confidently asserted explanations? There must be *something* that is shaping our practice and it is this something that we need to identify if we are to understand how folk psychology is deployed in practice and the goals we are trying to achieve when using folk psychology.

To begin to understand what could be happening when we construct explanations and predictions of behaviour, we'll explore the role that both social and cultural considerations play in shaping the explanations and predictions we provide. In particular, we'll review empirical evidence that suggests that one's folk psychology is significantly shaped by both one's social and cultural environment. This research shows that folk psychology is not a universally shared framework that is applied the same way in all cases by all people and that different social and cultural environments set different standards for the production and consumption of folk psychological explanations and predictions of behaviour.

If this is true, then our folk psychological practices are beholden to a host of cultural and social considerations that, importantly, need not be aligned with the characterization of folk psychology that we explored in Chapter Two. More specifically, the cultural and social considerations that shape our folk psychological practice need not be tied up with or limited by the goal of providing explanations and predictions of behaviour that are quasi-scientific in nature. A full discussion and argument regarding the implications of these empirical findings will be advanced in Chapter Six. For now I raise these suggestions to put the evidence below in context and prepare the reader for the forthcoming argument.

4.2 – Psychological Biases

The psychological research literature is filled with studies that explore psychological biases that affect our decision-making or the way in which we utilize and deploy information. More often than not the point of this research is to highlight some strange fact about human psychology that has implications for how we make decisions or use and assess information. The relationship between this research and folk psychology may not be obvious at first, but I think there are relevant links between how we make decisions and process information and how we make judgments about others and their behaviours, that warrants an investigation of this research as we investigate folk psychology.

First, traditionally construed, folk psychology just is a tool for helping us to explain and predict how people do or will think, reason, and decide. Any empirical evidence regarding how it is that we actually go about doing these activities has bearing on our investigation of a practice that has the aim of explaining and predicting these activities.[106]

Second, and more importantly for my purposes in this chapter and in particular this section, folk psychology is itself a decision-making and reasoning process. When we engage in folk psychology we input information and make decisions about how to ascribe or not ascribe beliefs, desires, intentions etc. to ourselves and others. How we recall, rely on, prioritize and more generally utilize information in decision-making will impact the way we make decisions as we attempt to explain and predict behaviour. If this research reveals something about our psychological make-up that would also apply to our folk psychological practices, then we have

[106] This harks back to a point I made in Fn. 15 of Chapter Two. There I noted that our contemporary notions of folk psychology can be traced back to and through the work of Aristotle, Hume, and Davidson. These philosophers and their work in the philosophy of action (how we reason and decide to act) has significantly shaped our contemporary notions of folk psychology. As such, an investigation into how we actually reason and decide is relevant to our discussion.

reason to investigate it. More specifically and as we will see, the quirks of reasoning and decision-making highlighted below help to demonstrate that perhaps we are not as good at folk psychology as we have been assumed to be and that perhaps we are not using folk psychology in a way that is consistent with the principles or expectations of a scientific endeavour.[107] At a minimum, I will argue, these studies hint at the suggestion that when we engage in folk psychological practices, the information we use in our practices may be less than a complete story[108] and this suggestion will set the stage for the empirical research explored in the following sections.

So how exactly is information relevant for decision-making recalled and deployed? In particular, is all information recalled with equal ease and do we deploy all information from different sources with the same carefulness? As it turns out, the answer to these questions appears to be no. A number of studies have shown that as we face opportunities to deploy information we tend to rely on the information that is most available to us on the assumption that this is something we have experienced most often.[109] However, it turns out, this information is often not that which we have experienced most often. This observation has become known as the availability heuristic or availability bias and it can play out in a number of ways. Often what is most available to recall is that which is most salient or that which leaves the greatest impact on

[107] There is a sense as well, as to whether or not the view of folk psychology discussed in Chapter Two is psychologically possible or likely. It could be that our cognitive system is built in such a way that the robustly explanatory and predictive characterization of folk psychology does not actually accord with the way that our cognitive system recalls and deploys information. This is not to say that these biases cannot be overcome, just that our default is not necessarily as is assumed.

[108] This has implications whether or not, for example, the mechanisms that underlie our folk psychological competencies are theory based or simulation based. If we attribute mental states on the basis of a theoretical competence, then our practice will be affected by how we process the information that feeds into our theoretical competence. Alternatively, if we attribute mental states to others on the basis of a simulation of our own decision-making, our simulation will be influenced by any quirks in our own information processing. In both cases, how we utilize or recall information can significantly impact our practices, and as discussed below, these tendencies may give rise to errors in our attribution practices.

[109] A fairly intuitive and natural suggestion. Information that we're most able to recall could presumably be information that we hear or read most often. This frequency increases familiarity and thus recall.

us; not necessarily that which is experienced most frequently. Similarly, the vividness of the experience also impacts our ability to recall it; a strong experience is just easier to remember and so is more readily available to recall. In contrast, experiences that are less impactful tend to fade to the background and become more difficult to recall.

Tversky and Kahneman (1973) helped to reveal this bias through a great experiment asking participants to determine what is more common in the English language, words that begin with the letter *r* or words that have *r* as the third letter. This is a challenging task, but the strategy any of us would deploy to answer it is clear.[110] We would simply start thinking about words that start with *r* and those that have *r* as the third letter. I'm sure many reading this have already started to try coming up with words that start with *r* and those that have *r* as the third letter. No doubt, those who have tried have noticed that thinking of words that start with the letter *r* while prone to instances of mental blocks, is not nearly as difficult as thinking of words whose third letter is *r*. As such, you may conclude that words starting with *r* are more common than those that have *r* as the third letter since they are more readily available to you and thus, must be something you experience with more frequency. In fact, Tverksy and Kahneman's research confirmed that people tend to reason in just this way.[111] Unfortunately, the availability heuristic has steered us in the wrong direction since words starting with the letter *r* are actually less common than words with *r* as the third letter (Tversky and Kahneman, 1973, pp.210-212).

Another way in which this bias can manifest is with respect to assessments of quantitative and qualitative data. Let's say you're purchasing a new TV. You've reviewed a host of data

[110] At least Tversky & Kahnemann present this as the most obvious route to the solution. I admit that I too explored this solution when first presented with the test.
[111] While not everyone concluded that words starting with the letter *r* were more common than words with *r* as the third letter, a significant majority of the participants did come to this conclusion (Tversky & Kahneman, 1973, p.212).

online and have looked through *Consumers Report* to access fairly unbiased and objective assessments of each TV you're considering.[112] Now suppose your neighbour catches wind that you're shopping for a new TV and gives a horrendous review of her own experience with one of the TVs you're thinking of purchasing. Which information source do you think will sway your decision more? As it turns out, we are systematically prone to lean towards the qualitative information because it is simply more salient. For example, a study conducted by Borgida and Nisbett (1977) demonstrated that students who reviewed evaluations of university courses consolidated from hundreds of class evaluations but also heard from a small panel of students were more influenced by the small handful of voices than the more reliable quantitative data. These handful of voices are simply more salient or vivid and so influenced the decision-making process far more than the quantitative data in spite of the obvious difference in reliability between the two data points.[113, 114]

These results demonstrate something general about the way in which we utilize information, but have implications for how folk psychology is deployed in practice. If we are ascribing mental states to others and ourselves, this practice will be significantly influenced by the way we recall information. Suppose that we've heard from a number of people that John is a really great guy who is regularly doing nice and generous things for people (i.e. a quantitative data set). But then we encounter John for the first time and he's short and rude with us (i.e. qualitative data point). Which information source is going to play a larger role in how we ascribe

[112] Adapted from Nisbett, Borgida, Crandall, & Reed (1976) who investigate the same type of story with respect to cars and Schwartz (2004) who recounts this research.
[113] I saw this play out all the time in my experience as a public opinion and program evaluation researcher. While I specialized in quantitative research, my clients would often want to add qualitative questions to their research or simply engage in qualitative research projects. For the most part, this research actually provided us with the least amount of information, but the information it did provide was salient and impactful. One comment from a stakeholder could often overshadow a few data points captured from multiple respondents.
[114] For a more detailed discussion of the ways in which we are susceptible to the availability heuristic in social situations see Nisbett & Ross (1980).

mental states to John in order to explain or predict his behaviour at some later point? If the availability heuristic is right, chances are that our own personal experience (although unique) may play a larger role in how we come to understand John's behaviour on that day because it is most salient and most vivid. Worse, explanations and predictions made on the basis of this more salient and vivid data point, may differ significantly from those generated on the basis of the larger data set and risk being less accurate as a result[115] giving us reason to wonder whether we are as good at explaining and predicting behaviour as has been assumed.

Not only, however, is information recall biased, it appears that how questions or decisions are framed can significantly change the way we select and use information to provide an answer or make a decision. Consider the following scenario. Parents, set for divorce, are fighting for sole-custody of their only child. Parent A is of average income, average health, works an average number of hours, has a reasonable rapport with the child and has a relatively stable social life. Parent B on the other hand, has an above-average income, minor health problems, travels a lot for work, has a very close relationship with the child and an extremely active social life. Who should we award custody to?

A study by Shafir, Simonson and Tversky (1993) showed that nearly two-thirds (64%) of respondents would award custody to Parent B (p.16). While Parent A was average in many ways, Parent B had two very positive features which must have been zeroed in on as relevant factors during the decision-making process and these factors outweighed the three negative features also associated with this parent.[116]

[115] The ideas being hinted at here re-appear in Section 4.3 below when I speak to the research conducted by Knobe & Malle (2002) which explores how negative or positive perceptions of a person can significantly alter the types of explanations we provide for their behaviour.

[116] This is an interesting observation on the way in which information is selected in itself, but the point I'm more interested in follows shortly.

Suppose the question was reframed not in terms of who should be awarded sole-custody, but who should be denied sole-custody. It turns out that answers to this question are quite different than the positively valenced version. In fact, under this scenario just over one-half (55%) would have awarded custody to Parent B, a significant decline in support (Shafir et al., 1993, p.16). The point of this experiment is to show that when faced with a decision we often latch on to whatever reasons we can find that will help us make our case and that this can change depending on how the question is framed. But the implications of this study are not limited to just the case at hand.

While these studies do not specifically address our folk psychological practices of producing explanations and predictions, or attributing folk psychological concepts to ourselves or others, the findings are pertinent to these practices. In particular, if how a question is framed impacts how we supply an answer, we can expect to see the same effect in folk psychology. That is, predictions of how a person will act could vary significantly depending on how the question is posed to the predictor. If folk psychology is subject to framing effects, it suggests that the folk can be divergent in their production of explanations and predictions and that someone[117] will likely be wrong in their explanation or prediction. This again, should lead us to wonder whether or not we are as good at explaining and predicting behaviour as has been assumed by the tradition.

It turns out as well, that simply being required to articulate the reasons for a decision can have an impact on how decisions are made. A study conducted by Wilson et al. (1993) asked college students to choose between five different posters, some comical and some works of fine art. Results of the study suggested that on the whole students preferred the works of art to the

[117] And in some instances, perhaps everyone.

more comical posters, unless, they were asked to explicitly write down what they liked or did not like about each poster (Wilson et al., 1993, p.334). In this scenario participants tended to favour and take home the more comical posters over the fine works of art when compared to those who did not have to write out reasons for favouring or not any of the posters. But what is perhaps most telling, respondents who had to write down their reasons were less satisfied with their choice one week later (Wilson et al., 1993, p.334).

The hypothesis here is that when forced to write down reasons for liking or not liking something, we may have a hard time articulating our precise preferences or reasons for the decision. It's not, I'm sure we will all admit, always easy to really know why we choose as we do. But, being forced to provide some reasons in writing leads us to grasp at those reasons that are most readily available to us and that appear to adequately explain our decision. It is easier, we might expect, to find these reasons for the funny posters since they are more tangible: it made me smile, the joke was funny, I liked the image, etc. In contrast, it is much more difficult to assess the reasons for the pleasure we find in great pieces of artwork. There is often simply an intangible element of the work that draws us in.

The results of this study suggest that being forced to articulate the reasons for an action can change the decision-making process. Again this may have implications for how folk psychology is used by the folk in their day-to-day lives. First, the tendency to latch on to whatever reasons we can when making a decision would apply just the same when providing an explanation for a behaviour. We'll latch on to whatever reasons are available to us that will help us provide a reasonable explanation, just as we do when we need to make a reasonable decision and explain this to someone. This is most obviously the case when explaining or predicting our own behaviour, since it mirrors the decision-making of the college students above and is also

evidently the case if we are simulating someone else's behaviour by trying to imagine ourselves in their shoes. However, we'll also likely butt up against the difficulty of articulating reasons when trying to ascribe mental states to others and so may rely on those beliefs and desires we can think of, rather than those that really caused the behaviour. This suggests first that our ability to explain and predict behaviour may be limited by our ability to identify and articulate reasons for behaviour, in other words, these results have implications for how successful we are at explaining and predicting behaviour. But these results also suggest that when explaining behaviour, the explanation we produce may be shaped by the act of justifying the practice itself. This task sets standards that we then set out[118] to achieve when providing an explanation. This relates to the discussion below in Section 4.3 and so I leave a more comprehensive treatment of this suggestion to that section.

Turning our attention back to information processing, the confirmation bias identifies a tendency in people to look for confirmation of our own viewpoint while simultaneously rejecting alternatives of equal merit. One study conducted by Lord, Ross and Lepper (1979) identified this bias by testing the impact of two articles with differing conclusions regarding capital punishment. Participants in this study were equally divided among those who supported capital punishment and those who did not. The group was presented with two articles recounting fake studies of equal merit and that were ultimately identical in every respect except the conclusion. One article argued in favour of capital punishment while the other argued against the practice. On the surface, reviewing two equally plausible articles recounting studies with contrary conclusions should leave a reader no better off than they were prior to reading the articles in

[118] I say "set out", but this practice is likely happening subconsciously or at an implicit level. Although it is likely that sometimes we explicitly attempt to create explanations and predictions that satisfy some standard, it is likely that the vast majority of our practices happen implicitly.

terms of being able to justify their own position. Of course, this is not what was found. In fact, participants in the study tended to find fault where there was none and discount the evidence of the article rejecting their own view. Moreover, and perhaps more importantly, people left the study feeling *more* confident in their own view than they had prior to the experience. In the face of effectively no evidence, they found evidence to confirm and support their own view and rejected equally plausible evidence regarding the contrary view.

The confirmation bias is similarly applicable to our folk psychological practices. In particular when faced with two alternative explanations or predictions the folk might be more inclined, if the confirmation bias applies, to pick the explanation or prediction that conforms with their knowledge of the actor[119] or that favours their instinctive assumption about the reasons for the behaviour or about what the person will do in the future. Again, this can lead to divergent explanations and predictions that may be inaccurate and should lead us to question whether or not our folk psychological competence is as sophisticated and accurate as is assumed.

Together these psychological biases raise the possibility that perhaps we are not as good at folk psychology as it has been assumed that we are.[120] As it turns out, decision-making and information processing are actually fairly messy businesses and there are a number of ways in which the biases explored above could steer us away from providing explanations and predictions of behaviour that are accurate, consistent across time frames, consistent across people, and generated on the basis of careful and thoughtful analysis. Moreover, these biases and the resulting implications for our folk psychological practice also cast a shadow of doubt on the

[119] As with above, this suggestion may in fact be borne out in the work of Knobe & Malle (2002), discussed below.
[120] The reader may have noticed a distinct absence in this discussion. More specifically, I have not discussed the fundamental attribution error and how we tend to explain behaviour on the basis of character traits and not situational or contextual factors. This absence is deliberate as the fundamental attribution error will be used as a central example in my discussion of the cultural factors that shape our folk psychological practices in Section 4.4 below.

idea that when we engage in folk psychological practices that we are doing so in a way that is consistent or appropriately aligned with a quasi-scientific account of folk psychology.

This is not to claim that there must always be clean overlap between the folk psychological and the scientific method (as noted in Chapter Two), but just to suggest that casting folk psychological explanations and predictions as being quasi-scientific or taking an even stronger position and saying that the folk are operating like scientists, may be harder to accept in light of all of the biases above and the negative effect they may have on the accuracy of our folk psychological practices. Additionally, this is not to say that these biases are insurmountable and that the folk are necessarily operating in a way that is different in kind from scientists, just to say that the difference in degree might be fairly significant leading to a practice that is a significant departure from where we'd like to be or where it has been assumed we are. Moreover, this is just the beginning of the discussion. While I think this foray into the psychological bias literature reveals something peculiar about our decision-making and information processing that has implications for our folk psychological practice, the main argument is yet to come. Next, we'll look at the related psychological literature on confabulation to find out what exactly the folk psychological stories we're telling ourselves and others are doing for us.

4.3 – Folk Psychological Confabulation

Telling stories is a central part of the folk psychological practice. We often tell elaborate stories that capture a variety of mental states and their interactions in order to explain or predict the behaviour of ourselves and those around us. The role of storytelling in folk psychology is so strong that is has led to some accounts of folk psychology that are entirely based on the storytelling or narrative experience. For example, Hutto (2008) goes to great lengths to make the

case that folk psychological concepts are not attributed by the use of a theory or by simulating another's experience in oneself. Instead, he proposes that folk psychology is simply a narrative practice. In this one respect, Hutto is absolutely correct; our explanations and predictions are embedded in often quite elaborate stories. Consider a favourite example of mine and others in the field.[121]

In the movie *The Princess Bride* the "genius" Vizzini goes to great lengths to out think The Man in Black (a.k.a. Westley) in a battle of wits. The challenge? To determine which of two wine glasses has had the deadly icocane powder dissolved into it and then drink from the other. In a memorable scene Vizzini employs folk psychology to its comical lengths in order to assess whether Westley has chosen to place the poisoned glass in front of himself or in front of Vizzini.[122] Alas, despite his claim that one should "Never go against a Sicilian when death is on the line", Vizzini dies as he failed to consider the possibility that Westley has developed a tolerance to iocane and has poisoned both cups.

I recall this story not just because Vizzini failed in his prediction, although this helps to serve a general point running through this chapter, but because Vizzini goes to such great lengths to reason through and predict Westley's behaviour. The comedic genius in this moment is the balance the writers strike between the elaborate nature of Vizzini's prediction, the familiarity of the structure of prediction he made to the audience, and the ultimate failure of his prediction. Together these factors lead to a laugh as Vizzini falls off his stool at the end of the scene. This example sets the stage for what follows. As we are about to see, we often construct explanations

[121] Kristin Andrews (2012) also makes reference to this example.
[122] In the interest of time I refer the reader to review the exchange in detail here: http://www.imdb.com/title/tt0093779/quotes?item=qt0482733

and predictions of behaviour that are elaborate, but these elaborate stories are often entirely mistaken.

In the late 1970's two prominent researchers, Richard Nisbett and Timothy Wilson, conducted a systematic review of the confabulation literature as well as a host of their own studies. As a result of this research, they concluded that when we provide folk psychological explanations and predictions of behaviour, we rely on an implicit set of "a priori"[123] causal relationships which express our cultural or social group commitments (Nisbett and Wilson, 1977, p.248). In a sense, when we utilize folk psychological concepts in practice to explain or predict some behaviour, we are engaging in a social enterprise that is in many ways shaped by a pre-established set of cultural guidelines or rules for engagement. What is most important to recognize about these rules of engagement, however, is that they need not necessarily produce folk psychologists that are scientists.[124] In other words, the social enterprise and the scientific enterprise need not be necessarily aligned and analogously, the rules we are abiding by may not be concerned with providing explanations and predictions of behaviour that accurately capture the causes of that behaviour. Instead, these rules may be pushing us to produce accounts of behaviour that are in accordance with a different standard, one set by our culture or social group. While the confabulation literature is plentiful, a quick exploration of two studies reviewed by Nisbett and Wilson will help to uncover the motivation for this account of the folk psychological practice.

[123] This is their term and I will not use it moving forward. Specifically, I find the label 'a priori' questionable in this case. Instead, moving forward I will simply examine the main claim, that our folk psychology responds to social and cultural considerations.

[124] Of course they could, but the point here is to just highlight the *possibility* that they do not. A more robust discussion about the relationship of these social and cultural considerations with the suggestion that we're operating as scientists will be undertaken in Chapter Six.

4.3.1 – Monkeys Swinging Across Rivers

Consider an experiment originally conducted in 1931 by the psychologist Maier. This study was designed to test our ability to report on our higher cognitive processes while we engage in the act of problem solving (Nisbett and Wilson, 1977, pp.240-241). Participants were placed in a laboratory where they found two cords hanging from the ceiling and an array of objects such as poles, extension cords and clamps lying around the room. The participants were then instructed to tie the ends of the two cords together. Unfortunately, participants soon discovered that the cords were placed far enough apart in the room that you could not hold on to one and still reach the other. Using the various props found around the room, participants could easily find three of the four available solutions. Generally participants could find these three solutions with some ease, but they often had some difficulty figuring out the fourth and final solution on their own. So, after a sufficient amount of time had passed the experimenter, who would be casually walking around the room, would gently swing one of the two ropes setting it into motion. This simple cue was enough to produce a reaction in all subjects, and within 45 seconds they quickly tied a weight to the end of a cord, set it in motion, grabbed the other cord and waited for the pendulum to come back their way. After catching the swinging rope they would tie the two ends together and complete the task. Immediately after completing the task, participants were asked how they came to this solution, effectively being asked to construct an explanation of their own behaviour.

Responses to this query generally came in the form, “It just dawned on me”, “It was the last solution available”, or “I just realized that a weight would let me swing the rope”. A professor of psychology, however, gave a much more elaborate response, suggesting that, “Having exhausted everything else, the next thing was to swing it. I thought of the situation of

swinging across a river. I had imagery of monkeys swinging from trees. This imagery appeared simultaneously with the solution. The idea appeared complete." (Nisbett and Wilson, 1977, p.241) Some explanations were rather banal, but some went to great lengths to capture the process that led to performing the task.

Although explanations were provided, none of the subjects could pick out the likely real cause of their behaviour, namely, the cue initiated by the experimenter when he gently swung the cord. In response to these results, Nisbett and Wilson conclude that the subjects were simply unaware[125] of the true causal explanation of their behaviour and yet would offer explanations anyway.

4.3.2 – Nylon Preferences

Consider another experiment. Nisbett and Wilson conducted a study where participants were approached in a commercial establishment under the guise of a consumer survey. The participants were asked to evaluate four different garments, as well as four identical nylon stockings choosing their favourite. In both cases, participants did choose a favourite, even when they were choosing between options that were identical. However, when analyzing the data an interesting result appeared. The analysis showed a strong correlation between the position of the item and how heavily it was favoured amongst the participants. According to the investigation, the right-most item was favoured the most. In fact, in the case of the identical nylons, the

[125] Shortly we'll review some literature discussing cultural variations in attention and the impact this has on the type and structure of folk psychological explanations. These results prompted me to consider that those in this early experiment were effectively falling victim to a blindness in their attention. As attention varies by culture, I reached out to Nisbett to assess the value of redoing this early experiment but with those of different cultures to explore whether other cultures are better suited to identify and mention this cause. In our exchange it was determined that there was merit to this study. Unfortunately, to my knowledge, no one has conducted this experiment and due to financial and logistical limitations I was not able to conduct it myself. It would be interesting to see the differences in explanation provided by people of different cultures in this experiment or similar experimental set-ups, and it would also be interesting to see if differences emerge between the explanations that actors themselves provide and explanations provided by external observers watching the action unfold.

correlation between favourability and position was four times stronger for the right-most versus left-most package. This is worth emphasizing. The position of the item, a rather irrelevant factor, was highly correlated with the level of preference reported by the participants. This is true when the options are identical but even when there were differences between the items that could encourage or detract from ratings of favourability.

For each decision, participants were asked to report on their choice and provide reasons for choosing which of the items was their favourite. No one ever mentioned the position of the article as a factor in their choosing. This is not surprising. We are not trained to or generally expected to provide explanations *like this*. In other words, we are not trained or generally expected to recognize an item's position in a line up as being relevant to our decision-making process regarding assessments of preference. That said, it is telling that this factor was never picked up on by the participants as it suggests that we may not be particularly well adapted to *always* provide explanations of our behaviour that are causally accurate since we are not well adapted to identify causes such as the position of an item. Even more telling, however, if it was suggested that the position of the item might have caused or influenced their decision, the participants seemed puzzled and unsure if the question was legitimate (Nisbett and Wilson, 1977, pp.243-244). In effect, their rejection of this cause suggests that their theory (for lack of a better term) of their own decision-making does not include a role for environmental factors such as the position of an item in relation to themselves.

There is no doubt, given the data collected, that the position of the item was a significant factor in determining the choice that was made. Yet, no subject was able to pick out this factor and incorporate it into their explanation (Nisbett and Wilson, 1977, p.243). What's more important to the story, however, is that explanations were provided and defended. Even more

profoundly, when choosing between identical items, reasons for the choosing were entirely fabricated, presumably attributing factors to one of the items that was simply not unique to it alone.

4.3.3 – Why we Tell Stories, and Don't Tell the Truth

In both of the cases just explored, participants gave a reason for their action, but these reasons did not accurately depict the state of affairs that led to their behaviour.[126] These studies once again suggest that we might not be as good at folk psychology as has been assumed, at least when it comes to explaining our own behaviour.[127] But more importantly, the results found that not only are we not sometimes in a position to develop an accurate account of our behaviour, but that when presented with the factors that lead us to behave as we did, we sometimes reject the presented reason as being foolish. Suggesting that the explanations we provide might also be designed in such a way as to fulfill some other standard, a standard that rejects the role of environmental factors in the production of behaviour.

[126] One might comment that the participants in these studies just simply didn't know or perhaps couldn't really know the cause of their behaviour. For example, one might suggest that introspection is quite difficult and that it is not a particularly reliable guide to the mechanisms that underlie our cognition. Peter Carruthers (2011) in his book *The Opacity of Mind*, makes an argument that is relevant to this point. More specifically, he argues that our introspective abilities regarding non-sensory states (e.g. beliefs, intentions etc.) are not transparent and are accessed in the same ways we access these states in others. Put another way and in the language of this section, our ability to access the reasons for our actions are interpreted in the same way we access the reasons for action of others. Without delving into the discussion in tremendous detail, this observation may help the argument unfolding in this chapter in two ways. First, it helps to support the claim that we are not particularly good at explaining our own behaviour by acknowledging that we may not have privileged access to the reasons for our behaviour. Second, it means that if we fail to explain our own behaviour particularly well, we may fail to explain the behaviour of others in the same ways and vice versa.

[127] The contexts within which participants in these studies have been required to explain behaviour are simultaneously strange and commonplace. They are obviously artificial constructs that are somewhat peculiar in nature, requiring the participants to identify an unusual cause of their behaviour. They are, however, commonplace insofar as the task is fairly benign; solve a problem and pick a favourite item. These are everyday tasks that we are faced with on a fairly regular basis and the experimental framework simply reveals something about our ability to construct accurate explanations and predictions and the nature of the explanations and predictions we offer. Put another way, these findings may be hard to observe in the everyday context since everything appears normal, but once the context can be manipulated and becomes a bit more unusual it's easier to spot a general feature about our folk psychological practice as it becomes more salient.

In response to this literature, Nisbett and Wilson suggest that instead of attempting to provide causally accurate explanations, we (explainers of behaviour) rely on an implicit set of socially, culturally, and experientially defined rules that guide us in explanation production and that determine what counts as an appropriate explanation. They suggest that there are cultural rules that state the relationships between various stimuli and the behavioural responses that follow, that there are implicit theories about causal relationships that our culture adopts, that we form a set of theories from our own experience of behaviour and stimulus response, and finally, that we may provide a novel hypothesis based on an extrapolation of our own experience when confronted with similar sets of stimuli and behaviour (Nisbett and Wilson, 1977, p.248). Importantly, we can extrapolate from this analysis. In particular, while these heuristics may be shaped and/or defined by true causal relationships and an impetus to provide accurate explanations, they need not be. Instead, the rules that we are operating under may be open to significant and important social and cultural influences that steer us away from causal accuracy for a number of instrumentally valuable reasons. For example, instead of providing explanations that pick up on the causally relevant factors that really did produce the behaviour, perhaps we construct explanations that rationalize or normalize the behaviour. In this way, the explanations we provide are the types of explanations that we're expected to provide, whether they pick out the causally relevant factors or whether they are a complete fabrication. So long as they meet the "sniff" test of our cultural or social grouping they will be endorsed and our explanation will be consumed as appropriate and acceptable.[128]

This suggestion helps to explain what is happening in each of the cases above. Participants in the study are providing responses to the question of why they chose as they did

[128] This will be discussed further in Chapter Six as I review Andrews (2012) analysis of our folk psychological practices as it pertains to explanation.

that do not conform to what would be expected of an appropriate explanation (e.g. accurately capturing the reasons for their behaviour), but that does conform to what we would expect someone in these situations to say. They can be simple or elaborate, but the participants provide explanations that *sound right*.[129, 130] Moreover, this analysis makes sense of the reaction participants have to the suggestion that something as simple as the position of an item can influence their favourability assessments. We find this suggestion implausible and react accordingly because it is not a part of our social or cultural understanding of behaviour. It is not expected, nor necessarily accepted, that our behaviour could be determined by such irrelevant factors and our expectation that behaviour conform to a standard of rationality (more below) excludes this explanation as being plausible.[131] As such, we react with surprise and even alarm when presented with this possibility even though the causes of our behaviour do fall within the realm of what could plausibly be a cause of our behaviour.

The plausibility of this suggestion is consistent with more contemporary research as well. Confirming the malleability of the explanations we provide, Joshua Knobe and Bertram Malle (2002)[132] highlight the importance of the rationalizing bias in the explanations we provide. In particular, we like to think of ourselves as being free and not confined by external sources. As such, when we provide explanations of our own behaviour we do so in a way that enhances our

[129] A question regarding whether the participants believe the accounts they provide or whether others believe the accounts provided could be raised at this point. I think with respect to the former question, there is reason to think that people are producing these explanations in good faith and with confidence. With respect to the latter question, the answer is a bit more elusive. That said, if the rejection of environmental factors is done because it doesn't conform to a social norm about how behaviour is caused, I suspect that others would accept the explanations offered by those in the studies above because they conform with the norm.

[130] This is not to say, for example, that identifying the swinging rope as a cue for developing a solution to rope task wouldn't also sound right. This is just to say that the explanations provided sound right or perhaps right enough, given the context within which the explanation is being provided. Of course the true explanation for the nylon preferences does actually sound less right, given our social or cultural expectations regarding what constitutes an appropriate explanation in that context.

[131] At least in the social and cultural group of those involved in the study.

[132] References are to a pre-print version of the cited article.

apparent freedom (Knobe and Malle, 2002, p.10). Importantly, our inclination to appear rational will alter the explanations we give. For example, when explaining our choice to carry an umbrella with us on a rainy day we will tend to say something like "Because it's going to rain today", not, "I think it's going to rain today". The difference between the two explanations is subtle, but important. In the former explanation, our cognitive state is characterized as *knowledge,* not *mere belief* as it is in the latter example.[133] While both phrases can be used in a rational explanation of behaviour, choosing the former over the later emboldens our decision to choose and confirms our own rational precision.

Similarly, positive behaviours will elicit explanations that enhance our role as the agent responsible for the action. Specifically, we will provide an explanation that locates the causes of the behaviour within us, thereby taking responsibility for them. In contrast, negative behaviours that we do not want to take responsibility for or have associated with us can be explained by citing environmental or situational factors; things we have no responsibility for (Knobe and Malle, 2002, pp.11-12, 17). Consider the difference between the explanations "The bus was early" and "I slept in" when it comes to explaining why I was late for work. In the former, I push responsibility for my lateness onto the environment, whereas the latter focuses on a characteristic of me as the agent of action. Compare even "I slept in" with "My alarm didn't go off". Even here there is a subtle deflection of causal responsibility in the latter case, attributing the resulting behaviour as being caused by something outside of the agent even when the agent themselves could be plausibly construed as ultimately responsible for the alarm being appropriately set. Our

[133] One may also interpret the difference between these two statements as a difference between stating a fact that is external to and independent of the individual versus citing a cognitive state that warrants justification or further investigation to ensure accuracy. I'm open to this interpretation and do not think that it alters the claim of this section. However, the distinction I've highlighted is more in line with how Knobe and Malle would characterize this difference (see also Malle, 2004 for further discussion on the implications of these kinds of subtle changes in language as it pertains to how the folk explain behaviour).

willingness to shape our responses based on the perceived valence of a behaviour reveals that we can use our folk psychological competence to satisfy a variety of standards and that we can tailor our folk psychological explanations depending on the context and/or the standard we are or want to be evaluated against.

Similarly, the types of explanations offered by actors and observers can vary. This has been known since the Jones-Nisbett hypothesis was proposed forty years ago (Jones and Nisbett, 1972) but the specific differences have been empirically refined since. In particular, Knobe and Malle recount work they have done showing that actors are significantly more likely than observers to use reason based explanations. As an actor describing one's own behaviour we have a vested interest in making ourselves sound like the cause of the action (with the exception of some the examples above); citing explanations that are internal and directly associated with agency. In contrast, observers do not have the same obligation to cast a behaviour in this light and they are perfectly willing to think of someone else as having false beliefs or undesirable traits (Knobe and Malle, 2002, p.11). As such, observers tend to cite reasons less often than actors describing their own behaviour (Knobe and Malle, 2002, p.15), instead outlining personality traits or other internal processes as well as environmental or situations triggers. All of this is to say that the type of explanation proffered imparts a different degree of responsibility on the actor or can portray the behaviour in more or less of a rational light.

Consistency would demand that similar situations demand similar explanations but something as simple as the actor changing from a stranger to myself may change the explanation I give. This suggests that our explanations can change in nature depending on the particular context within which we offer it. Certainly some of these differences are due to informational access such that even if we were trying to provide accurate explanations that capture the causes

of our behaviours, we might simply not be in a position to do so. However, some of these differences suggest that when we provide explanations of behaviour we are being influenced by other considerations, in particular, social or cultural considerations that shape the way we provide explanations. In effect, we are engaging in a sort of evaluation of behaviour that is constrained by social and cultural norms and in particular, our own drive to be viewed in a particular way. Explanations allow us to examine one another and show off our 'rationality' but they also give us the ability to discuss and think about our actions in a 'sophisticated' manner that need not be tied up with the need to provide explanations that are causally accurate.

Consider, for example, the psychologist's explanation above in the rope tying experiment. He crafted an elaborate story to show how he came to the final solution of the problem. This story shows sophistication, depth and an apparent understanding of oneself and cognition and creates a sense of superiority in the agent. Further, I challenge the reader to consider their own explanations for a moment. How many times have you told an elaborate story when perhaps there was a simple reason for acting? Now consider the fact that you might not have acted for the reason you cited, or that there was an essential factor that you simply did not mention. Consider again those who chose the right-most package of nylons and the details they gave in their explanation. While participants might not have been in a position to say that they chose the package they did "Because it was the last package", they reject this as possible explanation and instead provide one that appears reasonable and satisfies socially defined purposes, including the demand that behaviour appear rational and justified. That is, the explanations they offer instead suggest that they are motivated to give a socially defined and accepted account of their behaviour.

At this point, the true influence of cultural and social considerations is just slightly better than a mere suggestion.[134] The confabulation research explored by Nisbett and Wilson helps to first, demonstrate that we may not be as good at explaining our own behaviour as is assumed, but second, and more importantly, their research helps to identify the influence of social and cultural considerations by examining our failures. Moreover, the work of Knobe and Malle suggests that there is a malleability to our folk psychological practices, by identifying some of the social influences our folk psychological practices are subject to and in particular how considerations of rationality shape how we characterize behaviour.

In effect, it appears as though there is a social game we are playing when we provide explanations of behaviour, and this game exerts a great deal of influence on how our explanations are produced and what are deemed acceptable by those to whom we are providing these explanations; our social and/or cultural community. In these contexts, it makes sense to suggest that our folk psychological practices with respect to explanation are being driven by social and cultural considerations. However, the research and studies explored thus far in this chapter have all been within one culture, a Western or North American culture. To really test whether folk psychological practices are shaped by cultural or social norms we should test this hypothesis directly by comparing different social and cultural groups. Fortunately, this evidence exists.

[134] This is particularly true of the role of cultural considerations, which is only offered as a possible explanation of our confabulation behaviour by Nisbett and Wilson.

4.4 – Folk Psychology, Cross-Cultural Style

In his book *The Geography of Thought*, Richard Nisbett (2003) explores the suggestion that culture influences the structure of our folk psychological explanations and predictions. As we saw in Chapter Three Nisbett's work focuses on identifying the systematic differences between Western and East Asian people when it comes to a number of basic cognitive processes including perception, attention, and memory. While we explored the impact of these systematic differences on philosophical intuitions[135] in Chapter Three, these differences also lead to significant and important differences in the way events and behaviours are explained and predicted. Again, while the examples are plenty, we will focus on just a small sub-set to make the point.

4.4.1 – Psychological or Situational Explanations

Consider the following story. A recently fired postal worker, whose appeal to get his job back has failed, walks into his old job carrying a gun and murders his supervisor, the person who handled his appeal, and several employees and bystanders before turning the gun on himself (Nisbett, 2003, pp.112-113). How might we explain the gunman's behaviour? As it turns out, the answer to this question depends very much on your culture. Nisbett recounts a research project that tests just this question and which found significant differences in the explanations that American and Chinese people give.

In 1991 a Chinese physics student lost an award competition at an American university and after a failed appeal he shot his adviser, the person handling his appeal, several others and himself. Explanations in a local newspaper in the U.S. focused on the gunman's psychological state, attitudes and problems (Nisbett, 2003, p.111). In contrast, Chinese newspapers discussing

[135] Most notably for knowledge attributions, but also for assessments of reference.

the same story cited more contextual or situational factors like the student's relationship with his adviser, pressures in Chinese society, and the availability of guns in the U.S. Similarly, after a real life version of the post office story detailed above, a similar content analysis was conducted and the same results were found. American newspapers tended to focus on psychological or mental causes, while Chinese newspapers identified situational factors in their analysis of the events.[136]

To test whether these are differences in reporting style or more general cultural differences a more explicit experimental set-up was developed. Specifically, explanations of the murders were gathered and presented to both American and Chinese college students who were asked to review the explanations and rate the relative importance of the cited causal factors in each explanation. As it turns out, assessments of importance among the college students mirrored the results found through the newspaper content analyses. Americans, for example, found explanations of the form, "he had a violent temper", "there was a sinister edge to his character" and so on as having a higher importance in causing the behaviour, whereas Chinese subjects would tended think factors such as, "the gunman was recently fired" and "the post office supervisor was his enemy" (Nisbett, 2003, p.113) were more important causal factors.

Perhaps even more profoundly, when presented with a situational cause and asked to consider whether the murder would have occurred if this situational factor was not present, there

[136] A critic of this analysis might argue that the study does not really highlight a cultural difference, but reflects a difference in explanatory practices based on whether an individual is a member of an in-group or not. There is merit to this suggestion. However, this does not do away with the general theme emerging in this Chapter, that there are social and cultural considerations that shape our folk psychological practices. Whether the person being described is a member of an in-group or not is precisely the sort of social or cultural consideration that I am after. If we have a vested social or cultural interest in characterizing behaviour in a particular way based on in-group membership, then we have demonstrated the social or cultural character (whether the in-group is social or culturally defined) of folk psychology. Regardless, the analysis of newspaper articles is merely the beginning of Nisbett's analysis and the strength of the cultural differences analysis he offers increases with the volume of examples he has produced.

were significant differences between the American and Chinese students. Specifically, Americans were less likely than their Chinese counterparts to say that the outcome could have been avoided. This is a startling result, but not surprising given the earlier results. If situational factors are not presumed to be important in causing a behaviour, but mental states are, then removing situational information from the story but not mental state information should not produce a difference in the predicted outcome. In contrast, if situational factors are considered to be of more importance when it comes to causing behaviour, then removing this trigger from the scenario is likely to change the outcome.

Most obviously, these results confirm the suggestion that our folk psychology is culturally dependent by identifying stark differences between two cultural groups. But this finding may have implications regarding the assumption that our folk psychology is extremely successful as well. More specifically, these results suggest that people of two cultures may provide significantly different predictions based on the same information and it is possible these differences in predictive practices could impact the accuracy of those predictions. In fact, as we'll see below these cultural differences do sometimes lead to varying levels of success.

4.4.2 – Attribution Errors

Our attributions of mental states are similarly divergent. That is, there are systematic differences in the ways Westerners and East Asians assign mental states to those around them. Consider the following experiment. College students were invited to read a speech or an essay written by a fellow student, but explicitly told prior to reading the work, that their fellow student (call them the "target") was told to uphold the view they took and defended. That is, the target was instructed specifically to write an essay or speech in favour of, for example, the legalization of marijuana whether or not this was their particular point of view on the matter. After reading the

essay students participating in the study were asked to indicate what they thought the author's actual view on the issue was.

The right answer to this question should be "I have no idea". Participants in this study were alerted to the fact that the author of the paper was told to uphold a particular view and so they have, in effect, no information on the view of the author. While we typically and non-controversially assume the view being argued for is a view held by the author, in this particular experiment the participants have new information that explicitly rejects this assumption. As such, when asked what the author's actual view on the issue is, the right conclusion is that there simply is not enough information to know the view of the author.

However, in the experiment just described both American and East Asian students attributed the view espoused in the paper to the author. That is, if the paper was in favour of the legalization of marijuana participants of both cultures tended to ascribe this view to the author even knowing beforehand that the author was instructed to uphold the view argued for (Nisbett, 2003, pp.124-125). This commits all the students involved to an error in attribution.

The fact that all students, regardless of culture, tend to commit this error shows that people may not be as good at using their folk psychological tools as has been assumed. Presumably the students went through some roundabout justification in their minds to conclude that the author really was arguing for their own position and in order to make this mental manipulation successful, students would have had to deploy their folk psychological skills. What this tells us is that there is a strong psychological bias in us such that when faced with a well-articulated and reasoned argument we cannot help but attribute the view to the author. We fail to recognize, even when told that the author may not hold the view, that this could be an illusion.[137]

[137] I cannot help but look to American politics and the views espoused by some of the richest and most powerful men and women in the world as they try to persuade voters to give them more power and money (I choose American

Despite obvious evidence to the contrary, the immediate experiential information is prioritized over the contextual information as we assess and account for the behaviour we are observing, which leads to an error in our ability to accurately attribute mental states to others. Most importantly, this failure cuts across cultures, with people from different cultures committing the same attribution error. In effect and continuing themes explored above, we're really not as good at folk psychology as we think we are. While highlighting yet another failure in our folk psychology is important, the point of this section is to explore cultural variation in our folk psychological practice.

So what happens if instead of merely being told that the target was instructed to uphold a particular view, the subject is also instructed to write an essay in support of a prescribed view? That is, the subject is instructed to write an essay in defense of a particular view and is told that the target has been given the same task of writing an essay in defense of a particular view. In short, the results of the experiment change dramatically. As with before, participants in this study[138] were asked to read a paper from another student who was instructed to write in support of a particular view, but they too were asked to construct their own essay under similar circumstances. Once again, after reading the target essay participants in the study were asked to identify the personal view of the target author. Oddly enough, the American students tend to continue to make the same error and say that the author was in fact defending her own view. In fact, Nisbett writes, "[t]heir dispositional inferences about others were as strong as if they had not themselves experienced exactly the target person's situation" (Nisbett, 2003, p.125)

politics over Canadian politics as I think the truth of this claim is perhaps more evident in our neighbours to the South, especially when considering the Republican party, even though it is true of many Canadian politicians as well). Despite sometimes obvious evidence to the contrary, these people speak with confidence and assurance that seems to put others at ease and accept that the view being presented is really representative of the person giving it.

[138] To be clear, this is a separate study from the original where participants were merely informed that the target author had been instructed to defend a particular view.

suggesting that this tendency to discount the situational or contextual information when ascribing beliefs to others is very strong. In contrast, the Korean participants in the study did not make the same attribution error and were, in fact, nearly flawless in their attribution efforts (Nisbett, 2003, p.125). This is the cultural variation we were looking for.

It would seem that the American participants are simply too focused on the behaviour directly in front of them, namely a well-written and persuasive essay. This, presumably, leads them to ascribe mental states to the speaker on the basis of this evidence alone in spite of the situational evidence available to them. They are effectively blind to this information or this information is discounted in a striking way, even when it is something they experience directly and personally.[139] In contrast, the performance of the Korean participants changes. Specifically, while Korean participants in the first iteration of the experiment also erred, in the second iteration the situational information crosses a salience threshold and becomes a factor worthy of consideration in their attribution efforts. While it is perhaps possible to construct an experimental set up for the American participants that ultimately makes the information about the target salient enough to factor into their attribution, this simply was not achieved in the experimental set-up detailed above and so we can infer that if this threshold exists at all, it will be significantly higher for this cultural group than Koreans and likely for East Asians more generally. As with before, these experiments demonstrate that our folk psychological practices may not be as successful as

[139] In Section 4.2 above I discussed the way in which information recall is dependent, in part, on the salience or vividness of the information. The experimental findings here are likely related to those explored above, but there are unique implications of these results as well. First, this study demonstrates that different cultures may recall (or attend to) information differently. This suggests that there is cultural variation in the psychological bias noted above. Second, in the case where the participants experienced the target's position first hand, it is clear that the attribution error happens in spite of the salience or vividness of the experience. That is, this firsthand experience could be assumed to be more salient or vivid than merely being told that the target was instructed to defend a particular view, but in spite of this salience or vividness American participants still err in their attribution.

has been assumed, but the key take away is that our folk psychological practices vary significantly between different cultures.

4.4.3 – Predicting Jim

While the main point of the above discussion is to look at attribution errors in general, there is a strong focus on after the fact or explanatory behaviour. However, cultural sensitivity to the salience of situational information also leads to differences in predictive practices. Consider the following story devised by Nisbett and his colleagues.

> You just met a new neighbour, Jim. As you and Jim are taking a walk in the neighbourhood, a well-dressed man approaches Jim and explains that his car is broken down and he needs to call a mechanic. Then with a somewhat embarrassed voice, the man asks Jim for a quarter to make a phone call. You find that Jim searches his pocket and, after finding a quarter, gives it to the man. On another day Jim is walking toward the bus stop to catch the bus to work. As he is walking, a teenager carrying some books approaches Jim and politely asks him if he can borrow a dollar for a bus ride, explaining that he forgot his wallet at home and needs to get a ride to school. (Nisbett, 2003, p.126)

To test the impact of situational information on predictions of behaviour, Nisbett and colleagues presented this story to both American and Korean participants, followed by two variants. In Scenario 1, Jim searches his pocket to find several dollars. In Scenario 2, after searching his pocket he finds that he only has enough money for his own bus fare. All participants were asked to predict how Jim will act in response to the teenager's request.

As it turns out, predictions of Jim's behaviour varied by culture. Korean respondents were more likely than their American counterparts to recognize that in the Scenario 2 Jim had insufficient coinage to pay both his fare and the teenagers. As such, Korean respondents predict that Jim will not give the teenager money (Nisbett, 2006, p.126). Surprisingly, American subjects are less likely to make this same assessment. Instead, the contextual factor that Jim does not have enough money appears to be overridden by the fact that Jim is a generous person, a trait that is

alluded to early in the story. It would appear that yet again, there are significant differences between American and Korean participants in this study, suggesting that our folk psychological practices vary greatly between cultures, leading to widely divergent predictions of a fairly simple behaviour.

The same results were replicated under six different scenarios each with similarly different situational manipulations. In all six cases there were significant differences between the Korean and American participants. As in the scenarios discussed above one variant included situational information that would discourage a particular behaviour, while the other variant included situational information that would encourage the behaviour. In each case that included situational information that discouraged the behaviour, the Korean participants were significantly less likely than the American participants to predict that the discouraged behaviour would occur. Similarly, if the encouraging situational information was included, they predicted accordingly, in effect, showing a sensitivity that simply was not present for the American participants and that illuminates strong cultural differences in our folk psychological predictions. In Nisbett's own words, "Westerners tend to assume that events are caused by the object and Asians are inclined to assign greater importance to the context" (Nisbett, 2003, p.127) and these differences in the assumed locus of causation lead to important differences in the types of predictions we provide.

4.4.4. – The Social and Cultural Character of Folk Psychology

The examples explored throughout this section all support the same conclusion. There are social and cultural considerations that shape how the folk deploy folk psychology in practice. At a minimum, this means that in some contexts or in some cases folk psychology is not a universally shared and uniformly applied framework. Instead, depending on our cultural upbringing and the social demands of the situation, the way in which we use and apply our folk psychological

competence will vary. What remains to be seen, however, is the relative strength of this influence. How strong of an influence does our culture or social upbringing play in shaping our folk psychological practice? It could be that fundamentally we really are all interested in explanations and predictions of behaviour that are motivated by the same set of science-like standards and that our culture only comes in to play at a much higher level, shaping the style or the aesthetic quality of the explanations we give. Alternatively, it could be that the role of our culture is very significant and can be traced throughout the entirety of our folk psychological discourse.

The results just explored suggest that at a minimum the influence of our culture is more than a purely superficial one. But there may be limits to the influence our culture has on the folk psychological practice, most evidently demonstrated, in the "essay" study discussed in Section 4.4.2 above where both Westerners and East Asians commit the same basic attribution error. However, additional research regarding the development of these cultural differences helps to confirm that our culture has a profound impact on our folk psychological practices.

It turns out that children of different cultures may often be quite similar in how they use folk psychology and that differences do not begin to develop until adolescence (Nisbett, 2003 p.115). Research has been conducted to look at the differences in explanation producing behaviour between Americans and Hindu East Indians tracking these differences from childhood through to adulthood (see Miller, 1984). Results suggest that children in the two cultures do not really differ in the sorts of explanations they provide, but that as they get older and become a more complete member of their cultural community, their explanations begin to change. In particular, around adolescence explanations begin to diverge in significant and marked ways. At around 8 years of age, American and Hindu East Indian participants provided explanations citing

contextual factors (e.g. social norms, situation specific aspects etc.) or dispositional information (e.g. personality, attitudes etc.)[140] at about the same frequency, but as participants' age increased, the researchers found that the frequency with which dispositional information was cited by the American participants increased substantially, while the Hindu East Indian participants did not demonstrate the same effect to the same degree (Miller,1984, p.967).[141] In essence, this research suggests that a child must learn to give the explanations of their culture and learn to attend to the factors and phenomena that these explanations demand, and this takes time.

In Nisbett's words, "[i]f one thinks something is causally important one is likely to attend to it. So a cycle gets established whereby theories about causality and focus of attention reinforce one another" (Nisbett, 2003, p.114). It takes time, however, to become a full participant in this cycle. It takes time for the child to integrate into the culture and learn how to give the explanations that are expected, but once this has started, the cycle grows and develops, changing the way the child attends to, and explains the world of behaviour. In effect, our culture creates an expectation regarding how we ought to attend to the world and how we ought to understand behaviour. As if this is not enough evidence, the same researchers investigated Westernized Indians and found that their attribution practices were midway between Hindu Indians and Americans (Miller, 1984, pp.970-1). These results strongly suggest that our culture really does exert a strong force on how we attend to and perceive behaviour, as well as how we practice folk psychology.

In addition to the cultural standards that our practice appears to be shaped to meet, there is at least preliminary evidence that different cultures cut the world into different kinds when it

[140] For a full list see Miller (1984).

[141] It is interesting to note that both cultures started off the same, and that the American culture adopted the dispositional perspective more significantly. In particular, the same information or perspective that led to the failures noted above regarding the fundamental attribution error is also the perspective that is later to develop.

comes to behaviour. Some anthropologists (see for example Lillard, 1998) have suggested that different cultures use different concepts or terms to demarcate the world of behaviour. For example, Western cultures tend to have terms for the mind and the body that are distinct, whereas some East Asian cultures, namely the Japanese culture, do not split the "self" into these neatly divided objects (Lillard, 1998, p.12). Further, Westerners have a gamut of terms and ideas for the explanation of mental phenomena whereas some cultures may only have a single concept for that span of notions and often cultures do not even share the same concepts. For example, it is not clear whether Samoans have a robust concept of intention (Lillard, 1998, p.13), while this concept, for those of us in the West, is a core folk psychological concept that we use in a variety of explanatory fashions and that we divide into a multitude of more nuanced terms (e.g. intentional action, intention, intended).[142]

These differences in attention, assumptions about the locus of causation, the explanations and predictions we give, and even the entities we use combine to suggest that when we engage in folk psychological practices we are significantly limited by and shaped to conform to cultural norms or expectations. The suggestion emerging from this section is that when we engage in folk psychological practices, we are engaging in a socially and culturally defined endeavour that has a host of norms guiding our behaviour. These norms are not present just at the level of what factors should be cited in folk psychological explanations and predictions, but also in our how we attend to behaviour and how we attribute mental states to one another.

Importantly, these considerations significantly shape these practices and the result is dramatic differences in how our folk psychology operates across cultures. In effect, it is not too far of a stretch to say that different cultures have different folk psychologies. While there may be

[142] It is worth noting as well, that in Chapter Five we will explore this particular concept and how it is used by the folk in great detail.

overlap, it is not clear that there is a universally shared and practiced folk psychology. Moreover, these cultural considerations need not be shaped by the goal of providing accurate explanations and predictions of behaviour and can be quite distinct from this practice, suggesting that perhaps there are a variety of considerations beyond the scope of quasi-scientific explanation and prediction that influence the way we construct, use and digest folk psychological stories.

Demonstrating this cultural sensitivity has been the primary goal of this section, but it is worth noting that these cultural differences and the wildly different explanations and predictions they produce can in principle, and in some cases has already been shown to, lead to different explanations and predictions with varying degrees of success.[143] Our folk psychological practice, then, is in the business of producing accounts of behaviour that conform to our social-cultural discourse and this discourse need not be concerned with providing causally accurate accounts of behaviour and in fact, sometimes just isn't.

4.5 – Conclusion

It is apparent that we navigate the social domain with some skill. But just how good are we at providing explanations and predictions of behaviour? Throughout this chapter we explored numerous empirical studies that all suggest we are not as good at folk psychology as we think we are. There are strong psychological biases that impact the way we process and utilize information and that can have profound impacts on our own decision-making process as well as assessments

[143] One might think that a culturally defined folk psychology might be particularly apt at explaining and predicting behaviour of one's own culture and that failures likely only arise when using this culturally defined practice to assess behaviour of people from other cultures. I think this suggestion has merit and is consistent with the suggestion in Chapter Six that folk psychology is not a primarily explanatory and predictive framework but that it is, at its core, a regulative framework. However, in the context of this chapter there are two comments to make in reply and in defence of my assertion that cultural divergence helps us to defend the claim that we are not as good at folk psychology as has been assumed. First, cultural differences gives rise to an in principle possibility of failure: if two people give (wildly) different explanations or predictions of the same past or future behaviour it is likely that at least one of them is or will be wrong. Second, the cultural differences have already lead to observed differences in success. More specifically, a Westernized folk psychology with its focus on the agent and not situational factors is significantly more prone to committing the fundamental attribution error than the folk psychology of East Asia.

of how others make decisions and our attribution practices. But we also discovered that there are a number of social and cultural considerations that influence the way in which we construct, use and digest folk psychological stories. In particular, our explanations and predictions describe ourselves and others in rational ways or in ways that normalize our behaviour, our attitudes towards the actor change the way we explain or predict their behaviour, and there are important attentional differences between cultures that lead to markedly different folk psychological practices with varying levels of success. Importantly, none of these additional considerations need to be concerned with providing accurate explanations and prediction of behaviour. They could be, but they need not be.

The fact that we are not as good at folk psychology as is assumed should encourage us to pause, take a skeptical position for a moment and evaluate whether the traditional construal of folk psychology is right. That is, whether folk psychology really is so good it disappears or whether the adaptive utility of folk psychology really stems from its remarkable predictive success. Additionally, the fact that different social and cultural considerations appear to influence our practice in significant ways should also encourage us to pause, take a skeptical position for a moment, and evaluate whether the traditional construal of folk psychology as a quasi-scientific enterprise is accurate or whether at a minimum there are a number of considerations, either in addition to or replacement of, making accurate explanations and predictions that shape and drive our folk psychological practice. The questions these empirical findings raise will be explored in more detail in Chapter Six as we're not yet done kicking up the dust.

Thus far we've looked at the folk psychological practice at a fairly high level but we have an opportunity to narrow our focus and examine the layperson's use of specific folk psychological concepts to see how the practice is shaped at this more narrow level. Without

giving the punch line away, it looks like the way we use and attribute individual folk psychological concepts to one another is also significantly influenced by considerations not anticipated by the traditional account of folk psychology.

Chapter Five: Folk Psychology and Morality

5.1 – Introduction

In this chapter we turn our attention to the operation of a number of specific folk psychological concepts. By narrowing our focus from higher-level discussions of the structure and nature of folk explanations and predictions down to specific folk psychological concepts we can really investigate the folk psychological practice in detail. In particular, we can determine exactly how folk psychological concepts are used and attributed by the folk and begin to isolate and examine some of the key influences that drive these practices. While the last chapter included a brief review of some of our attribution practices, in this chapter we'll conduct a systematic review of a number of folk psychological concepts to help support a general picture about the entire folk psychological framework.

The aim of this chapter is to highlight a general finding that emerges with respect to how a number of folk psychological concepts are attributed in practice. More specifically, empirical investigations happening under the guise of experimental philosophy have revealed that whether a behaviour or the outcome of a behaviour is judged to be moral or immoral significantly changes the way that the folk attribute a number of folk psychological concepts to the agent of the behaviour. Put another way, moral judgements appear to shape the way in which the folk use folk psychology in practice by influencing the ways in which folk psychological concepts are attributed to others. In fact, in what follows we'll find that moral judgments significantly shape the way the folk psychological concepts of intentional action, causal responsibility, knowledge, belief, desire, and a host of other concepts are used in practice by the folk. These results suggest that the practice of folk psychology is intimately and significantly linked to the practice of morality.

The implications of this observation will be discussed in greater detail in Chapter Six, but if these results are indicative of a general effect whereby folk psychological practices are influenced by our moral judgements then we have even more reason to be skeptical of the traditional characterization of folk psychology that we explored in Chapter Two. Not only does our folk psychology appear beholden to a host of cultural and social considerations (Chapter Four), but the way we use folk psychology also appears to be subject to moral considerations. Again, none of these considerations need to be consistently aligned with the goal of providing quasi-scientific explanations and predictions of behaviour, and in many cases these findings might actually be at odds with this purpose. After all, from the traditional standpoint whether someone intended a result, believed their behaviour would cause an event, desired a particular outcome, etc. should not depend on the moral status of their behaviour. I raise these suggestions and concerns now to put the evidence below in context, but the primary purpose of this chapter is simply to explore and recount an array of studies conducted by experimental philosophers interested in folk psychological concepts.

5.2 – Folk Psychological Concepts

Traditionally, as we saw in Chapter Two, folk psychology is portrayed as involving the attribution of propositional attitudes in order to explain and predict behaviour. In particular, there has been a focus on the propositional attitudes of belief and desire, but the class of concepts is often reasonably extended to include concepts such as hopes, wishes, intentions, etc. In what follows, I will not be so restrictive. Among the propositional attitudes that are central to our folk psychology, I take a wide net that extends well beyond just the core concepts of belief and desire. Moreover, I will not restrict the class of folk psychological concepts to only those that take propositions as their content. Instead, properly construed, the class of folk psychological

concepts should include any concept that is productively used as a part of the practice and used to produce accounts of behaviour. In other words, and as we'll see below, if a concept is regularly and productively used in order to understand, explain, predict, and describe behaviour, if it is essential to the practice, or if it has relevance to how other more central concepts are deployed, then it should be included in an analysis of folk psychology and will be throughout this chapter.[144]

That said, I concede that the more traditional concepts of belief and desire and other propositional attitudes might form what we can call the "core" of the class of folk psychological concepts. While I'm willing to make this concession I think it is important to recognize that these concepts may not be as core to the practice as previously thought. That is, the fact that they are viewed as a core class of concepts may simply be as a result of the important theoretical and historical role they have played in our discussions of folk psychology thus far, and in the philosophy of action.[145] However, empirical research looking into how people choose to explain and predict behaviour suggests there is a wide array of concepts available to us when we engage in these practices and that these concepts need not be propositional in nature.

---

[144] A typical example of a concept that is productively used in folk explanations but that is not propositional in nature is personality traits. Plausible explanations for why Jane gave the homeless person her spare change could be "Because she is nice" or "Because she is generous". These personality traits do not have the same structure as propositional attitudes and yet form explanations that are familiar and regularly used by the folk. I grant that there may be a relationship between a trait and a corresponding propositional attitude (see for example Malle, 2004) but the point here is that traits typically fall outside the class of concepts included in folk psychology, and yet are productively used or at least have a strong relationship to the class of concepts typically included. Similarly, we might appeal to some statistical norm about human behaviour and offer an explanation such as "People generally stop at stop lights" when explaining why some stranger came to a stop at a red light. Again, statistical norms are not propositional attitudes and yet the explanation is sensible, familiar and one that the folk are likely to employ regularly and productively. While again I grant that there could be a propositional attitude that is related to the norm being employed, this relationship does not negate the point being made here. The concept being employed is not itself typically included in the class of concepts associated with folk psychology and yet has a role to play in the production of explanations, predictions and descriptions of behaviour. Throughout the chapter I will identify the relationship of the concepts in question to the typical class of concepts associated with folk psychology and offer reason as to why it should be included in the class of concepts (see as well Andrews, 2003).

[145] See relevant discussions throughout Chapter Two and in particular Section 2.3 – Folk Psychology as Explanation and Prediction.

Bertram Malle (see 2004 for an overview) has conducted a number of studies to better understand our explanatory and predictive practices and has developed a sophisticated catalogue to capture the different types of explanations we provide. First, Malle recognizes that the folk distinguish between intentional and unintentional behaviours and then have a variety of modes of explanation available to them to explain these behaviours. For intentional behaviours he argues that we can appeal to reasons (e.g. mental states on which the intention to act is formed), causal history of reasons (e.g. factors that lay in the background of an action's reasons), and enabling factors (e.g. how it was possible for the intention to become action).[146] For unintentional behaviours, Malle thinks we only cite "mechanical" causes that led to the behaviour without subjective reasoning or rationality and that mirror the types of explanations we provide for other physical events. Only the first category captures the propositional attitudes that are the focus of the traditional account (i.e. beliefs and desires[147]), whereas the remaining categories of explanation will not include explicit reference to propositional attitudes.[148] Others too have recognized that explanations do not always include reference to one's reasons (e.g. beliefs, desires, and other propositional attitudes) for action (see for example Andrews, 2003). Once we accept that beliefs and desires are not the only concepts that can be used to explain and predict behaviour, we can begin to explore just how some of these concepts function.[149]

[146] See Malle (2004) Chapter 4.
[147] Malle allows for the category of propositional attitudes to be broader than usual, explicitly endorsing the inclusion of the concept "valuings" in this category.
[148] It's worth noting that Malle believes that reason explanations are the default mode of explanation. In other words, this is the mode that the folk will employ unless there are reasons for doing otherwise (e.g. they lack the knowledge).
[149] Some may object at this point claiming that I have not properly defined the class of folk psychological concepts. I admit that I have not, but do not see this as a problem. First, it is an empirical question whether a particular concept is regularly and productively used and consumed as a part of our folk psychological practice and this project is simply beyond the scope of this project. Second, some of the concepts explored in what follows are undeniably a part of the folk psychological framework. That is, whatever this class of concepts looks like, the majority of those explored below will almost certainly fall into this class. Where there is some question regarding the concepts status are properly belonging to folk psychology, I have raised this below and addressed the concern directly. With these

5.3 – The *Folk's* Folk Psychology

Importantly, and unlike the preceding chapter, much of the empirical work we will explore here has been conducted by Experimental Philosophers. Leaving behind their comfortable but practical armchairs, philosophers of mind have entered the field of empirical research to determine how folk psychological concepts function in practice. Fortunately[150] for the practice, this endeavour has garnered a great deal of attention and a number of psychologists have joined in the investigation as well, and over time both the methodology and analysis of these endeavours have improved. As a result, we now have fairly robust empirical results regarding a number of folk psychological concepts.[151]

This research program started with a number of studies into how the concept of intentional action[152] was used by the folk, but as we will see many concepts have now been investigated empirically in order to properly and precisely understand how they function in practice. What we find through this investigation, however, comes as a surprise if the traditional account of folk psychology is right. In the previous chapter I spoke to the influence of social and cultural considerations on our folk psychological practice, in this chapter we will explore in detail the impact of moral considerations on how folk psychological concepts are used and attributed in practice. That is, we will reveal that our folk psychological practices are responsive to and guided by moral considerations.

---

considerations in place, I think it is fair to proceed without a clearly defined list of those concepts that belong in the class of folk psychological concepts.

[150] I say 'fortunately' as having psychologists join the discussion has brought with it a great deal of credibility and has helped to improve the methodology of the practice. This all has positive implications for the value of the research conducted and helps to ensure that the results are as accurate and meaningful as possible.

[151] Throughout much of what follows I will provide details regarding the participants of the study (i.e. how they were recruited), to signal the quality or representativeness of the sample as this directly relates to the inferences we can draw regarding the general population.

[152] Joshua Knobe deserves a significant amount of the credit for sparking this movement and investigation of folk psychology through empirical means and in particular, the concept of intentional action.

5.3.1 – Intentional Action

Much of the empirical research that has developed in the past decade concerning folk psychological concepts began with an investigation of the concept of intentional action and the phenomenon now known as the Knobe-effect, or, the side-effect effect. With a suspicion that the folk utilize this concept differently than philosophers have assumed, Joshua Knobe (see 2006b for an overview) employed empirical research methods to determine how exactly the concept of intentional action is used in practice. This concept, as Knobe notes (Knobe, 2006b, p.204), is extremely important for folk psychology and is a concept that people regularly use to distinguish between different categories of behaviour and has implications for how people explain behaviour as noted above (see Malle, 2004).

To do this Knobe developed a series of vignettes or short stories that involved an agent acting in a way that would cause a foreseen side-effect. Keeping the basic story identical, Knobe constructed two vignettes, one that portrayed the side-effect as positive or morally praiseworthy and another that portrayed the side-effect as negative or morally blameworthy. This experimental construction gave Knobe the means to test the impact changes in the moral valence of an outcome have on claims that someone acted intentionally to cause this outcome. One set of vignettes, which has become known as the Chairman vignettes, has played a very significant role in the debate that emerged. Knobe presented the following vignette to study participants:[153]

> The vice-president of a company went to the chairman of the board and said, 'We were thinking of starting a new program. It will help us increase profits, but it will also harm the environment.'
>
> The chairman of the board answered, 'I don't care at all about harming the environment. I just want to make as much profit as I can. Let's start the program.
>
> They started the new program. Sure enough, the environment was harmed. (Knobe, 2006b, pp.205-6)

[153] Participants were recruited from a Manhattan park.

Immediately after reading this story Knobe presented respondents with the following question: "Did the Chairman of the board intentionally harm the environment?" He collected their responses on a 7 point Likert scale and conducted analyses of the data. The data analysis revealed that the majority[154] of the participants believed that the Chairman *did* intentionally harm the environment. That is, the foreseen side-effect of harming the environment was intentionally caused by the Chairman.

In the second version of the Chairman vignette, Knobe made a simple modification changing the foreseen side-effect from a negative to a positive; from harming the environment to helping it. Presumably, this change re-cast the outcome as a positive and potentially morally praiseworthy outcome. This vignette reads as follows:

> The vice-president of a company went to the chairman of the board and said, 'We were thinking of starting a new program. It will help us increase profits, but it will also help the environment.'
>
> The chairman of the board answered, 'I don't care at all about helping the environment. I just want to make as much profit as I can. Let's start the program.
>
> They started the new program. Sure enough, the environment was helped. (Knobe, 2006b, p.206)

Once again, immediately after presenting participants[155] with this story Knobe asked them the following question: "Did the Chairman of the board intentionally help the environment?" and had respondents record their answer on a 7 point Likert scale. Knobe's data analysis revealed that in this scenario, the majority of respondents were not willing to say that the Chairman intentionally helped the environment.[156] More importantly, the mean rating in the harm variant and the help variant were significantly different (Knobe 2003, p.192). This result is quite

[154] 82% of the respondents said the Chairman intentionally harmed the environment.
[155] A different set of participants than were used in the "harm" variant. More specifically, Knobe randomly assigned participants to either the "harm" or "help" variant.
[156] Only 23% of respondents said the Chairman intentionally helped the environment.

surprising given that the behaviours were effectively the same, apart from the outcome with respect to the environment.

One might think that the folk appear to be responding inconsistently in the two cases. That is, the Chairman acted exactly the same way in both cases, performing the exact same action, with the same attitude toward the side-effect his action would produce. Yet, in spite of these carefully constructed similarities, the judgments of intentionality differ significantly between the two vignettes. Knobe concludes on the basis of this evidence that the difference in the moral valence of the side-effect outcome has produced the differences in judgments of intentionality his study highlighted. This minor and carefully manipulated difference must be the source of the variance in folk judgments of intentionality as it is the only thing that has changed between the two variants of the Chairman vignette. That is, we can say that since the only difference between the two vignettes is the change in moral information, that this change led to the significant variation in responses to the two cases.

Initially, there was a tremendous response to these findings. Some worried that the results were limited in scope or that the difference in attribution could be explained away. There might also be a problem with the results due to the demographics of Knobe's respondents. As we saw in Chapter Three some demographic factors are associated with differences in philosophical intuitions, and these differences may play a role in judgements of intentionality as well. But more importantly, Knobe's studies involved only random passersby in a Manhattan park or undergraduate students, each of which represent limited populations in their own respect. As a result it is unclear whether these differences can be extended to the wider population. However, these results are not unique in any way and, as we will see below, have been replicated

repeatedly in a number of different settings and by a number of different researchers, and in some cases in ways that address some of these early misgivings.

Interestingly, the asymmetry observed in attributions of intentional action is consistent across a number of demographic characteristics as well. Recall in Chapter Three the finding that socio-economic status was found to be associated with different intuitions regarding philosophical concepts such as knowledge (see Section 3.3 – Experimental Philosophy). In that section I raised the worry that if philosophical intuitions varied by socio-economic status that what constitutes knowledge for one group will not be the same as another group. Importantly, this worry does not appear to arise in the context of the folk psychological concept of intentional action.

Young, Cushman, Adolphs, Tranel, and Hauser (2006)[157] examined whether gender, education level, religious affiliation, English as one's mother tongue, and a background in moral philosophy had any impact on the way intentional action was attributed in response to the Chairman vignette. Young et al. found that none of these characteristics had any effect on the appearance of the asymmetrical attributions of intention action, that is, the presence of the side-effect effect. Across the board and regardless of these demographic differences, respondents all attributed intentionality to the Chairman asymmetrically and consistent with Knobe's original discovery. This is worth emphasizing. The side-effect effect is so robust that it cuts across a variety of demographic characteristics including even training in moral philosophy suggesting that the effect being observed is very stable and that it is essential to the function of the folk psychological practice. It is also worth emphasizing that Young et al. utilized a public facing

[157] These results are cited from an earlier and unpublished version of the paper. These results were irrelevant to the primary objective of the paper and were presumably removed before the final paper was published. That said, this data appears to have been vetted through other scholars and has been included in other peer-reviewed and published papers, including, for example, Beebe & Buckwalter (2010).

website to invite participants ranging in age from 18 to 88 to participate in their study. Importantly, while this sample would still show a strong self-selection bias, this approach is a positive step to overcoming some of the sampling problems associated with Knobe's original approach.

Adding more fuel to the fire, Knobe and Burra (2006) utilized the same experimental set up as Knobe's original investigation, this time asking only Hindi-speaking respondents to participate.[158] They made a minor change to the question supplied after the vignette was read, using the Hindi word "jaan-bujhkar", which has roughly the same meaning as "intentionally". Their study found that the same pattern of asymmetrical responses appears again. This is a particularly interesting and strong result given the discussion in both Chapter Three and Chapter Four. Recall that in Chapter Three we found significant differences in the way Westerners or people from the Indian sub-continent attributed knowledge to an agent (see Section 3.3 – Experimental Philosophy). Similarly, in Chapter Four we found that folk psychological practices varied significantly depending on whether one is of Western, East Asian, or Hindu East Indian descent (see Section 4.4.4 – The Social-Cultural Character of Folk Psychology). In contrast to the cultural variation that was observed in these discussions, this study suggests that the concept of intentional action is attributed in the same way by members of different cultures that are likely to show variations with respect to other folk psychological practices and philosophical concepts.[159] Extending this further, we might suggest that these results indicate that moral considerations may drive our use of the concept of intentional action to a greater degree than the cultural differences that drive other observed differences in our folk psychological practice.

[158] It should be made clear that these participants were also English speaking and that the vignette was presented in English with the exception of the target phrase. Participants were recruited from South Asian clubs at Princeton University and Yale University.

[159] Of course, this is just one small study, but the result is suggestive and has been often endorsed by others.

In fact, the asymmetry embodied in the side-effect effect is so robust that it appears as early as four and five years old (Leslie, Knobe, and Cohen, 2006). This is an important result as well. At ages four and five, children are just starting to pass the false-belief task, a task that is supposed to indicate when a child can appropriately attribute the concept of belief to others. Yet at this early age where belief ascriptions are really just forming, children are already using the concepts related to intentional action in a way that mirrors adult users. That is, children of a very young age have already acquired the skills required to utilize these concepts, but in a way that indicates sensitivity to moral information and an ability to vary their ascriptions of the concept accordingly. In fact, this level of sensitivity to moral information suggests that the sophistication with which children are using the concept is quite high. Specifically, it would be easier to apply the concept in the same way in each case as the attribution process would involve one less judgment: a moral judgment.

The literature on experimental studies regarding intentional action is extensive, but this selection of studies has already demonstrated the point. The side-effect effect is a robust phenomenon that is present in a variety of experimental set-ups, across cultural and demographic differences and is entrenched from a very young age. In other words, there is overwhelming evidence that moral considerations influence the way the folk psychological concept of intentional action is employed by the folk in practice. As a result, we should raise a red flag regarding an interpretation of this concept such that our competence ought to be understood in explanatory and predictive terms. Instead, these results suggest that our folk psychological concept of intentional action might have a moral purpose. Put another way, and in Knobe's words, "the competencies underlying our folk-psychological concept of intentional action have

been shaped in a fundamental way by a very different sort of use" (Knobe, 2006b, p.205) than has been assumed by the traditional construal of folk psychology.

5.3.2 – Causal Responsibility

Empirical investigations into the way in which the folk attribute and use the folk psychological concept of causation have also been undertaken. More specifically, experimental attempts have been undertaken to understand how the folk assign causal responsibility to others. As it turns out, just as attributions of intentional action are sensitive to moral judgments, so too are attributions of causal responsibility.[160]

Starting as mere speculation, Knobe (2006a) claimed that moral judgments would play a direct role in folk attributions of causal responsibility. To probe intuitions and generate support for his speculation, Knobe designed the following hypothetical thought experiment:

> Lauren and Jane work for the same company. They each need to use a computer for work sometimes.
>
> Unfortunately, the computer isn't very powerful. If two people are logged on at the same time, it usually crashes. So the company decided to institute an official policy. It declared that Lauren would be the only one permitted to use the computer in the mornings and Jane would be the only one permitted to use the computer in the afternoons.
>
> As expected, Lauren logged on the computer the next day at 9:00 am. But Jane decided to disobey the official policy. She also logged on at 9:00 am. The computer crashed immediately. (Knobe, 2006a, p.68)

Importantly, the computer will not crash unless both Jane and Lauren jointly log on at the same time. Their combined behaviour is the cause of the outcome, and yet Knobe hypothesized that the folk will respond to this case in such a way as to attribute more causal responsibility to Jane than to Lauren. If Knobe's hunch is correct, that the folk really would attribute greater causal

[160] Much of the following discussion is inspired by the commentary offered in Roxborough & Cumby (2009) as well as an unpublished article that Jill Cumby, Ben Fraser and I produced.

responsibility to Jane than to Lauren, then we can reasonably assume that this difference in attribution must be caused by the only difference between Jane and Lauren's behaviour; the moral valence of their respective behaviours. Thinking of it counterfactually, had both been allowed to log onto the computer, the responsibility would be equally distributed.[161] While Knobe never tested this hypothesis himself, others were curious enough to investigate how the folk would attribute causal responsibility to Jane and Lauren in the above case.

To this end, Livengood, Sytsma and Rose (2011) found that Knobe's hunch was in fact correct. Participants in their study[162] were invited to consider the following, slightly revised, version of the Jane and Lauren case Knobe developed:

> Lauren and Jane both work for a company that uses a mainframe that can be accessed from terminals on different floors of its building. The mainframe has recently become unstable, so that if more than one person is logged in at the same time, the system crashes. Therefore, the company has instituted a temporary policy restricting the use of terminals so that two terminals are not used at the same time until the mainframe is repaired. The policy prohibits logging into the mainframe from the terminal on any floor except the ground floor.
>
> One day, Lauren logged into the mainframe on the authorized terminal on the ground floor at the exact same time that Jane logged into the mainframe on the unauthorized terminal on the second floor. Lauren and Jane were both unaware that the other was logging in. Sure enough, the system crashed. (Livengood et al., p.12)

Participants were then asked to state their agreement[163] with the following two probes:

- Lauren caused the system to crash.
- Jane caused the system to crash.

[161] Equal distribution does not entail that both Jane and Lauren get a midpoint rating on the scale. Instead, so long as the two get equal ratings at any point on the scale, their causal responsibility can be viewed as being equally distributed. In contrast, an unequal distribution would be one where there was a difference in the rating between the two actors.

[162] Participants were recruited through a public facing website, were native English speakers, 18 years of age or older, with at most a minimal training in philosophy. Minimal training in philosophy was taken to exclude philosophy majors, those who have completed a degree with a major in philosophy, and those who have taken graduate-level courses in philosophy. (Livengood et al., p.13)

[163] Respondents were asked to rate their agreement using a 7 point Likert scale with anchors 1 (Strongly Disagree), 4 (Neutral), 7 (Strongly Agree).

Confirming Knobe's speculation, participants in this study were significantly more likely to say that Jane caused the system to crash, rather than Lauren[164] (Livengood et al., p.13). As such, the participants' attributions of causal responsibility for the system crash demonstrated sensitivity to the moral difference between Lauren and Jane's behaviour. To repeat and emphasize this point, the asymmetry in judgments of causal responsibility that Livengood et al. found must be due to the only difference between Jane and Lauren's behaviour, the fact that one action was sanctioned or permissible while the other was not sanctioned and therefore, not permissible. This moral difference leads the folk to make different judgements regarding the causal responsibility of each actor.

Knobe originally characterized the difference between Jane and Lauren's behaviour as a moral one, and I have followed suit. With empirical verification of the asymmetry in causal responsibility attributions we can follow Knobe's lead and conclude that moral differences are influencing the way in which we use and attribute the folk psychological concept of causal responsibility. However, the difference between Jane and Lauren's behaviour can be characterized in another way. If this characterization is appropriate and successful, then we cannot claim that a moral difference is driving the asymmetry in attributions of causal responsibility since this other characterization might be driving the effect. That is, if there are multiple ways of describing the difference between their behaviours, then we cannot claim that judgments of causal responsibility are influenced by moral considerations.

Julia Driver makes this exact point (2008a, 2008b) and so argues that this particular vignette cannot demonstrate that moral information is the source of the asymmetry. Instead, she points out that we can interpret Jane's behaviour as also being 'out of norm' since, after all,

[164] Mean ratings of agreement were 2.42 for Lauren and 5.21 for Jane.

employees do not typically break company policies. In other words, Jane's behaviour is unusual or atypical. Thus, she proposes that *this* feature of Jane's behaviour, not the moral valence of her behaviour, is the real feature that explains this difference in the folk's responses. In effect, she claims that the folk "are more likely to make attributions of causation to events that do not conform to norms" (Driver, 2008b, p.459). Like Knobe's original suggestion this too is just mere speculation and Driver does not cite empirical evidence to support her claim. That said, Driver's suggestion is a good one and seriously challenges the legitimacy of Knobe's speculation and the claim that the moral difference between Jane and Lauren's behaviour explains the empirical results Livengood et al. collected.

In response to Driver's alternative analysis, Knobe paired up with Ben Fraser (2008) to develop a new vignette and conduct an experiment.[165] The structure of the case is identical to the Jane and Lauren vignette. That is, two actors jointly and equally contribute to an outcome but one actor is acting in a way that can be construed as a moral violation. Importantly, in this new vignette the two behaviours are both equally 'in the norm' or typical. By ensuring that both the moral and immoral behaviours are typical, Knobe and Fraser controlled for the influence of atypicality and so have ensured that if any differences in attributions of causal responsibility arise, this is not due to differences in typicality but the differences in the moral valence of each behaviour. Their vignette, which has become known as the Pen case, is as follows:

[165] Participants were recruited from an undergraduate philosophy class at University of North Carolina – Chapel Hill.

> The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take the pens, but faculty members are supposed to buy their own.
>
> The administrative assistants typically do take the pens. Unfortunately, so do the faculty members. The receptionist has repeatedly emailed them reminders that only administrative assistants are allowed to take the pens.
>
> On Monday morning, one of the administrative assistants encounters Professor Smith walking past the receptionist's desk. Both take pens. Later that day, the receptionist needs to take an important message… but she has a problem. There are no pens left on her desk. (Knobe and Fraser, 2008, p.443)

Immediately after reading the vignette participants were asked to rate their agreement[166] with each of the following claims:

- Professor Smith caused the problem.
- The Administrative Assistant caused the problem.

To be clear, both Professor Smith and the Administrative Assistant contributed equally to the outcome, they both must take a pen in order for the receptionist to be left with none, thereby causing her unfortunate predicament; neither actor acting alone would be sufficient to cause the outcome. Yet, when participants rated their agreement with each of the above claims, Knobe and Fraser found that participants were significantly more willing to say that Professor Smith caused the problem. Assessments of causal responsibility were significantly higher for Professor Smith than they were for the Administrative Assistant,[167] revealing an asymmetry in judgments of causal responsibility in response to the presented scenario (Knobe and Fraser, 2008, p.443).

Importantly, unlike in the Jane and Lauren vignette, in this vignette Professor Smith's behaviour was perfectly normal for him, it was something he would do all the time. As such, Driver's analysis of the Jane and Lauren asymmetry does not apply in this case. As a result, the

[166] Knobe and Fraser used a 7 point Likert scale with anchors -3 (Not at all), 0 (Somewhat), +3 (Fully).
[167] Mean ratings of agreement for Professor Smith and the Administrative Assistant probers were 2.2 and -1.2 respectively.

asymmetry in attributions must be due to the only difference between Professor Smith's and the Administrative Assistant's behaviour, the fact that the latter's is perfectly moral while the former's is a moral violation. Put another way, the two behaviours are descriptively the same, but normatively different and this normative difference must be leading to the observed asymmetry in the attributions being made. Therefore, we can re-conclude that attributions of causal responsibility appear to be sensitive to moral considerations.

In reply, Driver (2008b, p.459) clarifies her notion of being 'out of norm'. Importantly, she recognizes that an action can be out of norm in two ways; statistically and evaluatively. To be statistically out of norm simply requires a behaviour to be something one does not normally engage in doing, or that we typically do not observe others doing. For example, statistically or typically speaking, people do not walk backwards around the office. Alternatively, there is a sense in which something can be evaluatively out of norm as well. We make this judgment based on whether some behaviour deviates from a norm or ideal conception of action. For example, people do not pick their noses in public.

Immoral actions are often atypical in both senses. That is, people simply do not tend to perform immoral actions and when we see someone doing something wrong, we view it as being unusual in this statistical sense. But, immoral actions are also judged to be out of norm in the evaluative sense. Driver concedes that neither the Professor nor the Administrative Assistant acted statistically out of norm, but maintains that the professor did act evaluatively out of norm. To be clear, while immoral behaviours are often statistically atypical, Professor Smith was known to be a frequent violator of morality and so his behaviour was statistically typical. That said, it was still wrong and in this sense 'out of norm'. As such, Driver claims that Knobe and

Fraser's data is perfectly compatible with her analysis as the Professor's behaviour is still 'out of norm'.

In response to his debate, Jill Cumby and I (see Roxborough and Cumby, 2009) realized that by holding statistical atypicality constant Knobe and Fraser's data does not help us assess whether their analysis or Driver's more complex analysis is correct. More specifically, Driver recognizes that a behaviour can be out of norm in more than one way (i.e., statistically or evaluatively) and her analysis suggests that either can play a role in driving our attributions of causal responsibility. In effect, by controlling for statistical atypicality, Knobe and Fraser's experiment cannot assess the role that this information might play in assessments of causal responsibility. As such, the investigation into the role of each information type is not complete and so we developed a variant of the original Pen case to investigate whether statistical atypicality plays a role in folk causal attributions. By separating statistical from evaluative atypicality, we can assess whether and the extent to which each form of typicality informs folk attributions of causal responsibility. If both types of typicality influence folk attributions of causal responsibility then a complete account of the concept would have to take this into consideration, and so we should prefer Driver's analysis over Knobe's.

This analysis is actually a familiar one and harks back to Hart and Honoré's (1959) claim that "the notions of what is unusual and what is reprehensible by accepted standards both influence the use of causal language" (p.56). Here we see a clear commitment to the idea that both sorts of information play an important role in our use of causal language. Thus, we developed a modification of Knobe and Fraser's vignette to fully test for the role of statistical atypicality and to provide a nice comparison between our results and those of Knobe and Fraser. Our vignette is as follows:

> The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take the pens, but faculty members are supposed to buy their own.
>
> The administrative assistants typically do not take the pens. Unfortunately, the faculty members do. The receptionist has repeatedly emailed them reminders that only administrative assistants are allowed to take the pens.
>
> On Monday morning, one of the administrative assistants encounters Professor Smith walking past the receptionist's desk. Both take pens. Later that day, the receptionist needs to take an important message… but she has a problem. There are no pens left on her desk. (Roxborough and Cumby, 2009, p.209-10)

I suspect that the reader will have spotted the difference. In our vignette, we reintroduced statistical atypicality, but this time, as a feature of the Administrative Assistant's behaviour. That is, while Professor Smith's behaviour remained statistically typical, but evaluatively out of norm, we modified the Administrative Assistant's behaviour describing it as moral (evaluatively typical) but statistically atypical as she does not usually take pens from the receptionist. Like Knobe and Fraser we asked our participants[168] to rate their agreement to the following two claims using the same scale as Knobe and Fraser:

- Professor Smith caused the problem.
- The Administrative Assistant caused the problem.

As Knobe would predict, our participants were significantly more willing to attribute causal responsibility to Professor Smith than they were to the Administrative Assistant.[169] Thus, we confirmed that moral information (or in Driver's terminology evaluative information) plays a role in the folk attributions of causal responsibility. In fact, our results suggest that evaluative information has a far greater impact on judgments of causal responsibility than statistical information. That is, since making the Administrative Assistant's behaviour statistically atypical

[168] Participants were recruited from an undergraduate philosophy class at York University.
[169] Mean ratings of agreement for the Professor Smith and Administrative Assistant probes were 0.9 and -0.98 respectively.

did not flip the asymmetry in assessments of causal responsibility, we can conclude that the moral information, not statistical information, plays a more significant role in the evaluation of causal responsibility.

We did, however, notice an unexpected difference. Specifically, our change to the vignette led to a small but statistically significant change in willingness to attribute causal responsibility to Professor Smith. That is, agreement with the statement "Professor Smith caused the problem" was lower among our respondents than Knobe and Fraser's, and this difference was statistically significant. Thus, we argued that statistical atypicality does have an effect, just not as we would expect, and not to the degree with which we were hoping to find. Our results motivated us to accept Driver's more complex analysis of the concept, requiring that a model of how the folk attribute causal responsibility will need to include information regarding the statistical and evaluative atypicality of the behaviour, not just evaluative information as Knobe has proposed. And for a time, this became the consensus position, endorsed in principle by Driver, ourselves, and later by Hitchcock and Knobe (2009). However, the hallmark of empirical research is having replicable findings and a new study was conducted which suggests that our conclusion regarding the importance of both moral and statistical information is not correct.

Sytsma, Livengood, and Rose (2012) were not satisfied with the consensus view nor the experimental efforts of Knobe and Fraser and Roxborough and Cumby. As such, they set out to conduct a more complete set of studies that comprehensively analyzed folk attributions of causal responsibility. They began by articulating a difference between population-level statistical norms and agent-level statistical norms (Sytsma et al., 2012, p.816). That is, we can analyze a behaviour as being statistically typical or atypical in relation to a group they belong to (population-level) or in relation to the agent's own past behaviour (agent-level). Thus far, the

discussion of whether or not statistical atypicality shapes folk attributions of causal judgments has relied on population-level statistical norms. More specifically, in the Pen vignette utilized by Knobe and Fraser and again by Roxborough and Cumby, both Professor Smith's and the particular administrative assistant's behaviours are typical or atypical in so far as they belong to a particular group, faculty members or administrative assistants respectively. Systma et al. correctly note that there could be differences between how population-level and agent-level statistical norms shape folk attributions of causal responsibility, and so a further investigation is warranted. Additionally, Sytsma et al. had concerns regarding the empirical data relied on to date for various methodological reasons (Sytsma et al., 2012, p.817). As such, they were not convinced that population-level statistical atypicality in fact directly influences folk attributions of causal responsibility as the consensus view claimed.

In response, they develop their own view which they call the responsibility view. In contrast to the consensus view which claims that information about statistical norms and in particular, whether a behaviour is atypical, directly impacts folk attributions of causal responsibility, the responsibility view claims that statistical norm information in general may only sometimes impact folk attributions of causal responsibility and do so indirectly through our normative judgments regarding the responsibility of an agent (Sytsma et al., 2012, p.816). On the basis of this view they make two predictions. First, they predict that population level statistical information will not impact folk attributions of causal responsibility since they do not see this information as being particularly relevant to the question of who is normatively responsible for a problematic outcome. In contrast, this view does predict that statistical norm information will impact folk attributions of causal responsibility when it is agent-level and demonstrates that a behaviour is typical. They reason that this information is relevant to the

question of who is normatively responsible for a problematic outcome, which will then influence the attribution of causal responsibility. To make this point clear, they invite us to consider a habitual jaywalker and an occasional jaywalker who both cause accidents.[170] Their hypothesis is that the habitual jaywalker will be judged as more blameworthy because of their pattern of behaviour and this in turn would affect attributions of causal responsibility (Sytsma et al., 2012, p.816). Put another way, they expect patterns of behaviour to be a relevant factor when making assessments of responsibility since someone who regularly engages in a particular behaviour is more likely to know of the bad outcomes that could come from their behaviour and this normative judgment will influence how the folk attribute causal responsibility.

In their first set of studies, Sytsma et al. analyzed each of the possible permutations regarding the Pen vignette (Sytsma et al., 2012, p.817). In particular, they utilized Knobe & Fraser's original vignette, the Roxborough & Cumby variant, as well as one more vignette[171] structuring the Professor and Administrative Assistant scenario to match the Jane and Lauren scenario where the behaviour that is evaluatively atypical is also statistically atypical at the population-level. The last vignette is as follows:

[170] It's worth noting that both behaviours are impermissible and so evaluatively atypical.
[171] The final permutation where the administrative assistant's behaviour is permissible and atypical, while the Professor's behaviour is impermissible and atypical was not used.

> The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take the pens, but faculty members are supposed to buy their own.
>
> The administrative assistants typically do take the pens. In contrast, the faculty members typically do not take the pens.[172]
>
> On Monday morning, one of the administrative assistants encounters Professor Smith walking past the receptionist's desk. Both take pens. Later that day, the receptionist needs to take an important message…but she has a problem. There are no pens left on her desk. (Sytsma et al., 2012, Appendices p.1)

Armed with all three permutations, they invited participants[173] to consider the stories and respond to the following probes using a similar scale to the one used by Knobe and Fraser:[174]

- Professor Smith caused the problem.
- The Administrative Assistant caused the problem.

As the responsibility view predicts, results of this study suggest that population-level statistical norm information has no impact on the folk attributions of causal responsibility (Sytsma et al., 2012, p. 817). In fact, across all three permutations of the vignette the exact same result was observed. Agreement with the claim "Professor Smith caused the problem" was significantly higher than agreement with the claim "The Administrative Assistant caused the problem", and there were no statistically significant differences in the level of agreement between permutations (Sytsma et al., 2012, p.817-818). Given the lack of variation across the permutations, we can conclude that the evaluative difference (the impermissibility of the Professor's behaviour) is the source of the asymmetry observed in the assessments of causal responsibility for Professor Smith

[172] Sytsma et al. removed the phrase "The receptionist has repeatedly emailed them reminders that only administrative assistants are allowed to take the pens" from this variant, presumably because atypical behaviour on the part of the faculty members would not likely warrant this admonishment.
[173] Participants were recruited through a public facing website, were native English speakers, 18 years of age or older, with at most a minimal training in philosophy (Sytsma et al., 2012, p.817).
[174]The standard 7 point Likert scale ranges from 1 to 7 and this is what Sytsma et al. used. In contrast, Knobe & Fraser used a 7 point scale starting from -3 and going to +3. Roxborough & Cumby used the same scale to be consistent with Knobe & Fraser. While a change in scale can change results, Systma et al. consistently employed their preferred Likert scale for all permutations of the vignette to produce comparable results within their own experimental design.

and the administrative assistant. These results cast doubt on the role of population-level statistical norm information in folk attributions of causal responsibility and thus the consensus view, which carves out a direct role for this information. In contrast, these results confirm the prediction of the responsibility view which predicts that population-level information will not impact the attributive practice.

However, Sytsma et al. conducted a further study[175] removing permissibility information from the Pen vignette to more directly test the influence of population-level statistical norm information on folk attributions of causal responsibility.[176] As it turns out, there were no statistical differences between all permutations[177] of the vignette when the permissibility information was removed (Sytsma et al., 2012, p.817). If population-level statistical norm information was to have a direct impact on folk attributions of causal responsibility, this would be the time to observe it and yet there were no differences between the vignette permutations.

This is fairly damaging for the consensus view, and again confirms the prediction of the responsibility view as articulated by Sytsma et al. Moreover, this reinforces the claim that the asymmetry observed when permissibility information was included must be due to the moral information and not some combination of the moral and statistical information.

In order to support their second hypothesis that agent-level statistical norm information influences attributions of causal responsibility, Sytsma et al. conducted a further study. In this study they updated the Pen vignette to include agent-level statistical norm information as opposed to population-level statistical norm information and made reference to a particular

[175] The same sampling strategy was employed in this study, as was used in the previously discussed study.
[176] Sytsma et al. were concerned that perhaps the role that population-level statistical norm information could be playing a role in folk attributions of causal responsibility, but was just being overridden by the presence of the moral information.
[177] It is worth noting that Sytsma et al. included the fourth vignette where both behaviours were atypical in this study.

administrative assistant, John. Since the vignettes used mirror the structure of those described above, I include only one vignette[178] below to demonstrate the changes in language:

> The receptionist in the philosophy department keeps her desk stocked with pens. The administrative assistants are allowed to take the pens, but faculty members are supposed to buy their own.
>
> One of the administrative assistants, John, almost always takes a pen every time he passes the receptionist's desk. In contrast, one of the faculty members, Professor Smith, almost never takes a pen when he passes the receptionist's desk. The receptionist has repeatedly e-mailed reminders that only administrative assistants are allowed to take the pens in order to ensure that everyone is clear on the point.
>
> On Monday morning, John encounters Professor Smith walking past the receptionist's desk. Both take pens. Later that day, the receptionist needs to take an important message… but she has a problem. There are no pens left on her desk. (Sytsma et al., 2012, Appendices p. 5)

Again, for each permutation, participants[179] in the study were invited to consider a vignette and rate their agreement with each of the following probes:

- Professor Smith caused the problem.
- John caused the problem.

Recall that Sytsma et al. hypothesized that agent-level statistical information would indirectly influence folk attributions of causal responsibility and that in particular, typical behaviour would be more likely to be viewed as causally responsible. This is precisely what was observed in the above study as results showed that assignment to a particular vignette affected folk attributions of causal responsibility in statistically significant ways (Sytsma et al., 2012, p.817). Perhaps most importantly, they found that when permissibility and agent-level typicality are in competition, the statistical information "wins out" (Sytsma et al., 2012, p.817). Looking at the vignette where Professor Smith acts impermissibly but atypically for him, while John acts

[178] The vignette below is the impermissible/atypical and permissible/typical variant. For a complete list of the vignettes visit: http://philsci-archive.pitt.edu/5372/
[179] The same sampling strategy was employed in this study as those above.

permissibly and typically for him, folk attributions of causal responsibility regarding John are actually significantly higher than for Professor Smith[180] (Sytsma et al., 2012, p.817). This confirms their prediction that the folk are more likely to attribute causal responsibility to an agent when they behave in a way that is typical for them then when it is atypical for them. To further strengthen their case, they utilized the same vignettes with permissibility information removed to examine the role of just agent-level statistical information and again results confirmed their hypothesis (Sytsma et al., 2012, p.818). More specifically, when permissibility information is removed respondents are more likely to say that the agent who behaved agent-level typically caused the problem than when they behaved atypically. In the words of the responsibility view, the agent acting agent-level typically is judged as being more blameworthy than the agent acting atypically and so folk attributions of causal responsibility vary accordingly.

On the basis of this series of studies, Sytsma et al. conclude that the consensus view is incorrect. More specifically, they conclude that population-level statistical information does not directly affect folk attributions of causal responsibility. While this new development discounted some of my own research, their findings are actually more supportive of the general claim being developed in this chapter than my earlier research on this issue. In fact, Sytsma et al.'s findings in relation to the original Pen vignette and its variants more strongly support the claim that moral information regarding a behaviour directly influences how the folk attribute causal responsibility than the consensus view I had originally endorsed. However, these results also suggest that agent-level statistical norm information has a role to play in folk attributions of causal responsibility. Importantly, however, the responsibility view holds that the role of agent-level statistical information is subservient to the task of determining normative responsibility for an

[180] Mean ratings of agreement for Professor Smith and John are 2.85 and 4.81 respectively.

action and the results of these studies are consistent with this claim. In other words, agent-level statistical information is important not because it directly influences how the folk attribute causal responsibility, but because it helps the folk make a normative assessment about blameworthiness or responsibility which then feeds into how they attribute causal responsibility (Sytsma et al., 2012, p.820).[181]

Fiery Cushman (2010) conducted a series of experiments[182] that confirm the general findings recounted above, but with a different set of vignettes. He began by constructing three different scenarios all detailing an individual making a choice that led to a bad outcome. By manipulating the background information of the scenario Cushman created a morally good and morally bad variant of each of the three scenarios. That is, the background information was manipulated in such a way that the bad outcome was the result of a morally good or morally bad decision on the part of the actor in the scenario. What he wanted to know was whether or not determinations of the causal role of the actor would be significantly different between the morally good and morally bad variants despite the fact that the act and outcome were identical. Participants in his study were presented with both the morally good and morally bad variants of the same scenario and asked a series of questions after reading each scenario. The structure of the three scenarios is similar enough that I will only explicitly recount the morally good and morally bad variants of one scenario, which Cushman calls the "Boat Case" (Cushman, 2010, p.11). The morally good and morally bad variants of the Boat Case were as follows:

[181] It would have been helpful if Sytsma et al. had also asked respondents to explicitly make this normative judgement in the survey. This would help to confirm the precise role that the agent-level statistical information is having on the normative judgment.

[182] Participants voluntarily accessed the survey through a public facing website. If respondents took less than four seconds to review the scenario and begin answering the survey, they were excluded from the final survey sample. Participants were presented with both variants of all three scenarios (6 scenarios in total) and were randomly assigned to one of two presentation orders. The first 100 respondents to each presentation order were included in the final survey sample (Cushman, 2010, pp.8-9).

> **Morally Good Variant:** Justin is driving his motorboat in the bay when he notices some swimmers in trouble. There are five swimmers drowning at the end of a narrow channel in front of Justin. In between Justin and the drowning swimmers is another swimmer who is safe and not in trouble. If Justin takes the narrow channel to the five drowning swimmers and saves them, the wake from Justin's boat will wash over the safe swimmer, drowning him. If Justin does nothing, the five swimmers will drown and the one swimmer will remain safe. Justin knows exactly what the consequences of his decision will be. Wanting to save the five swimmers, Justin decides to take the narrow channel. The one swimmer drowns. (Cushman, 2010, Appendix p.1)
>
> **Morally Bad:** Dave is driving his motorboat in the bay when he notices some friends nearby. His friends are fishing at the end of a narrow channel in front of Dave. In between Dave and his friends is a swimmer who is safe and not in trouble. If Dave takes the narrow channel to visit his friends, the wake from Dave's boat will wash over the safe swimmer, drowning him. If Dave does nothing, he won't see his friends and the one swimmer will remain safe. Dave knows exactly what the consequences of his decision will be. Wanting to see his friends, Dave decides to take the narrow channel. The one swimmer drowns. (Cushman, 2010, Appendix p.1)

Notice that Justin and Dave behave in exactly the same way and the outcome or side effect of killing the one is the same in both cases. Their actions were deliberate and they had full knowledge of what they were doing. The only difference in the two cases is the background information regarding why they drove up the narrow channel. That is, whether they chose to save five people (morally good) or go see their friends (morally bad). After reviewing the scenario, participants in the study were asked questions relating to the causal responsibility of Justin and Dave as well as the moral valence of their respective behaviours.

For the causal attribution question, Cushman asked participants to evaluate "How much of a role did [agent's name] play in causing the one to die?" and provided a 7 point scale anchored from 1 "None at all" to 4 "Some" to 7 "Very Large" (Cushman, 2010, p.9). Similarly, for the moral judgment question was "From a moral standpoint, [agent's name]'s decision was:", with a different 7 point scale anchored from 1 "Obligatory" to 4 "Permissible" to 7 "Forbidden" (Cushman, 2010, p.9).

Consistent with the experiments discussed above, results of this study showed that participants were significantly less likely to say that Justin caused the one to die when compared to ratings for the causal role of Dave (Cushman, 2010, p.12). More specifically, the morally bad variant received significantly higher ratings than the morally good variant and given that the two scenarios are identical except in regards to the moral valence of the behaviour, we can conclude that this difference is driving the observed result, just as we saw with the Pen vignettes above. The same general result was in fact observed for the other two scenarios as well, suggesting a general trend across multiple cases (Cushman, 2010, p.12).

Cushman also found that participants viewed the two variants as being significantly different in terms of the moral valence of Justin and Dave. In fact, results confirmed that the Justin variant was considered significantly less morally bad (close to the permissible anchor on the scale) than the Dave vignette (close to the forbidden anchor on the scale) and similar results were found for the other two scenarios as well (Cushman, 2010, p.12). On the basis of these results Cushman ran an additional analysis to assess the correlation between the moral judgment and the causal judgment within each participant. The analysis did in fact reveal a significant correlation between these two judgements (Cushman, 2010, p.13), which further confirms the relationship between the two judgements and this was found for all three scenarios.

Cushman opted to conduct an additional study since the variance in the ratings Justin and Dave's behaviours received was generally low (Cushman, 2010, p.14). He hypothesized that the deliberate act of driving down the narrow channel may be viewed negatively regardless of whether it was done for morally good or morally bad reasons[183] (Cushman, 2010, pp.14-15) and

[183] He made this prediction on the basis of previous research he conducted, which suggested that omitting to save a life rather than bringing about its end would be viewed as less morally bad (Cushman, F. A., Young, L., & Hauser, M. D., 2006).

with greater variance in the results, a more robust result might be observed. For this additional experiment Cushman tweaked all three of the original scenarios to describe a case of omission as opposed to actively bringing about the one's death. The revised Boat Cases were as follows:

> **Morally Good Variant:** Justin is driving his motorboat in the bay when he notices five swimmers drowning. Justin starts towards the swimmers when he notices another swimmer in the opposite direction who is also drowning. If Justin turns around, he can save the one swimmer, but the five will drown. If Justin continues towards the five swimmers he can save the five, but the one will drown. Justin knows exactly what the consequences of his decision will be. Wanting to save the five swimmers, Justin decides to continue towards them. The one swimmer drowns. (Cushman, 2010, Appendix, p.2)

> **Morally Bad Variant:** Dave is driving his motorboat in the bay when he notices some friends who are fishing. Justin starts towards his friends when he notices a swimmer in the opposite direction who is drowning. If Justin turns around, he can save the one swimmer, but he won't be able to go fishing with his friends. If Justin continues towards his friends he can go fishing with them, but the swimmer will drown. Justin knows exactly what the consequences of his decision will be. Wanting to go fishing with his friends, Justin decides to continue towards them. The one swimmer drowns. (Cushman, 2010, Appendix, p.2)

Again, notice that Justin and Dave's actions and the outcomes are identical, Cushman simply manipulates the background information of the story to change the moral valence of the choice made by the actor. As before, participants were asked to assess the causal role that Justin and Dave played in killing the one and were asked to assess the moral valence of their decision.

Similar results to those in the first experiment were observed as participants rated Dave's causal role as significantly greater than Justin's causal role, despite both actions and outcomes being identical (Cushman, 2010, p.16). Likewise, participants rated Dave's behaviour as being significantly worse than Justin's (Cushman, 2010, p.16). These results were consistent across all three scenarios and Cushman found a significant correlation between moral judgments and causal judgments in all three scenarios (Cushman, 2010, p.17), clearly suggesting that in the case

of omission we still find the asymmetrical effect that has been observed even when the act is deliberate.

Given these two experiments Cushman concludes that although it is clear that the folk think that agents have considerable causal responsibility for bad side effects, this is significantly less the case when the consequence is brought about by a morally good choice, a result that is consistent with those detailed above (Cushman, 2010, p.17).

However, Cushman set out to try and answer an additional question. More specifically, he was interested in assessing the role moral information plays not just in our immediate attributions, but in the inferences we make about causality in general. Put another way, Cushman was interested in trying to develop some empirical evidence to assess whether moral judgments can affect the underlying representation of causation the folk have and use in inferential ways (Cushman, 2010, pp.18, 21).

To begin to assess this question empirically, Cushman modified the three omission scenarios and the corresponding morally good and morally bad variants to include mention of a prior causal contributor that had a role to play in bringing about the bad outcome (Cushman, 2010, p.21). The morally good and morally bad variants of the Boat Case in this experiment were as follows:

> **Morally Good Variant:** Justin is driving his motorboat in the bay when he notices five swimmers drowning. Justin starts towards the swimmers when he notices another swimmer in the opposite direction who is also drowning. The one swimmer was overcome by an unusually large wave. If Justin turns around, he can save the one swimmer, but the five will drown. If Justin continues towards the five swimmers he can save the five, but the one will drown. Justin knows exactly what the consequences of his decision will be. Wanting to save the five swimmers, Justin decides to continue towards them. The one swimmer drowns. (Cushman, 2010, Appendix, p.5)

> **Morally Bad Variant:** Dave is driving his motorboat in the bay when he notices some friends who are fishing. Justin starts towards his friends when he notices a swimmer in the opposite direction who is drowning. The one swimmer was overcome by an unusually large wave. If Justin turns around, he can save the one swimmer, but he won't be able to go fishing with his friends. If Justin continues towards his friends he can go fishing with them, but the swimmer will drown. Justin knows exactly what the consequences of his decision will be. Wanting to go fishing with his friends, Justin decides to continue towards them. The one swimmer drowns. (Cushman, 2010, Appendix, p.5)

Notice in both cases that an unusually large wave has caused the one to begin to drown prior to Justin or Dave being faced with a decision. As with before, the acts and outcomes are identical, but the background information has been manipulated to generate a difference in the moral valence of the two variants. Participants in this study were asked to assess the moral standing of Justin and Dave's decisions but were also asked "How much of a causal role did the tidal wave play in the death of the one?" using the same scale as was used to assess Justin and Dave's behaviour in the previous experiments (Cushman, 2010, p.21).

Again, and consistently across all scenarios, Cushman found that participants rated the morally bad variant as significantly worse than the morally good variant (Cushman, 2010, p.23). The more interesting and novel result of this experiment emerged when looking at the target question regarding the causal role assigned to the tidal wave. As it turns out, participants rated the causal role of the tidal wave significantly lower in the morally bad variant when compared to the morally good variant (Cushman, 2010, p.23). In other words, when Dave acts in a morally objectionable way, participants rated the causal role of the tidal wave significantly lower than in the case where Justin acts in a more morally acceptable way. Cushman argues that this result is consistent with the causal discounting view, which predicts that if we attribute a causal role to one event, that this licenses us to infer that other causal candidates have a lesser role in bringing about the event (Cushman, 2010, p.22; see also Morris and Larrick, 1995). But what is important

for our purposes here is, as Cushman notes, that "subjects not only attribute a greater causal role to morally bad agents, they also use this attribution to infer that other situation factors that contributed to the bad consequence played a lesser causal role." (Cushman, 2010, p.23) As such, we have evidence to suggest that moral judgments do not exert their influence on judgments of causal responsibility in a limited way, but instead these moral judgments influence the inferences we make regarding the "larger body of causal knowledge connected to the bad consequence" (Cushman, 2010, pp.22-23) as well. Additionally, although Cushman does not emphasize this observation, it appears as though our moral judgments do not just influence our assessments of the causal role of agents, but also situational factors. So while the Pen vignettes above and Cushman's first two studies demonstrated that the folk attribute causal responsibility to agents in a way that is sensitive to the moral valence of their behaviour, these results suggest that the moral valence of an agent's behaviour changes the way we attribute causation to non-agent entities as well.

Together, these experiments paint a fairly consistent picture of the way in which the folk use the folk psychological concept of causal responsibility. Specifically, they demonstrate that moral considerations influence the way in which this concept is employed by the folk in practice. As with intentional action, this should raise a red flag since these results suggest that our concept of causal responsibility might have a moral purpose or be driven by moral considerations which may very well be at odds with the traditional interpretation of folk psychology.

5.3.3 – Ascriptions of Knowledge

The relationship between knowledge and belief is a close one. In epistemology we talk about, for example, knowledge in terms of justified true belief, belief formed on the basis of a reliable process, or as a web of interconnected and supported beliefs. However we cash out knowledge, belief will play a fundamental role in understanding how we attribute knowledge. In fact, we might say that you cannot know something if you do not believe it.[184] Unfortunately, the converse is not true. The fact that you can believe something, even justifiably believe something, and not know it has played havoc with philosophical attempts to understand the concept of knowledge. In fact, this is precisely the sort of problem that thought experiments such as the famous Gettier case prey on. We can easily describe scenarios where someone believes something to be true or even justifiably believes something to be true that happens to be true, but are not in possession of knowledge. In the *Truetemp* case, for example, the actor clearly believes that he has correctly provided the ambient temperature, in fact he has. But does he know the temperature or is having an unknown implant in the brain that produces the thought, really knowledge producing? In the *Fake Barn* case, if the actor does not realize he is driving in an area of town where papier-mâché barns are common, can he really be said to know the barn is real? My point here is to merely highlight the close relationship between believing and knowing since belief has played such an important role in how we conceive of and understand folk

[184] I am open to the possibility that there may be limited cases where someone would rightly say "I know X, but I really don't believe X". While I have a hard time parsing this out philosophically, I am more open to the possibility that the folk might use the concepts in this way but I do not have the empirical evidence to assess this claim either way. I am, however, comfortable saying that knowledge at least generally entails belief or that when one knows something, more often than not they believe it as well. I will speak of the relationship between knowledge and belief in a way that implies certainty with respect to this issue, but I am open to curtailing the claim slightly. However, and most importantly, this has no impact on my general argument. While here and below I imply that the close relationship between knowledge and belief has implications for how the folk psychological concept of belief is used, if I must reframe this close relationship in less certain terms, than the implications I suggest with respect to belief can be similarly re-framed. The general point will not change, although the degree of impact may simply be somewhat more muted than this discussion implies.

psychology. In effect, because knowledge and belief have such a close relationship, and belief is a core folk psychological concept, there is good reason to think that knowledge is a core concept as well, just perhaps a slightly less famous one.

Given the central role belief has played in folk psychology and its close relationship with knowledge, I submit that knowledge is also a central folk psychological concept. In other words, by attributing knowledge to others (or ourselves) we can improve or assist our folk psychological practices. This should be clear from the fact that considering one's knowledge is paramount to determining what one believes and thus, providing information on a central concept of our folk framework.

In fact, the relationship of knowledge to folk psychology is clearly expressed by some of those who have responded to the original Chairman vignettes. In particular, much of the philosophical discussion that followed Knobe's presentation of the results invoked the concept of knowledge to understand the Chairman's behaviour and understand how the folk would view his behaviour. For example, Hugh McCann (2005) states that "in both vignettes, he [the Chairman] knows perfectly well what he is doing" (p.739) and Cushman and Mele (2008) echo the same sentiment when they say "One [Chairman] knowingly harms the environment by starting a profit-making venture and does not ''care at all'' about harming it; the other knowingly helps the environment by starting a profit-making venture and does not ''care at all'' about helping it." (p.172) Others focus more on the harm case where knowledge might play an important role. For example, Guglielmo and Malle (2010) write, "the CEO obviously knows that his action will bring about harm"[185] and Wright and Bengson (2009) propose, "In HARM the chairman presumably knew that his action would have a bad outcome." (p.27) Each of these cases

[185] This quote is from an unpublished version of the paper cited.

indicates that philosophers assume knowledge is being ascribed to the Chairman and that this information is factoring into the assessment of intentionality. In other words, by applying knowledge ascriptions to the Chairman we can better understand the reasoning process that the folk go through when they attribute the concept of intentional action. These quotes reveal a strong commitment to the role of the concept of knowledge in folk psychology, and in particular, in ascriptions of other folk psychological concepts.

But, as Beebe and Buckwalter (2010) point out, the vignette never actually discusses the Chairman's knowledge states; instead it only discusses what information is presented to the Chairman. As such, in each of the quotations above philosophers have effectively assumed how the folk will ascribe knowledge to the Chairman. We can do better than this by investigating these ascriptions empirically to determine the extent to which the folk ascribe knowledge to the Chairman in both vignette variants.

Inspired by the side-effect effect literature concerning intentional action, Beebe and Buckwalter investigated whether a similar effect could be found for the concept of knowledge. To conduct this investigation they utilized the original Chairman vignette, but instead of asking participants to ascribe intentionality, they asked respondents to ascribe knowledge to the Chairman. Specifically, participants[186] were presented with the original Chairman vignette, and asked to assess their level of agreement[187] regarding the chairman's knowledge that the new program would harm/help the environment.

As it turns out, they found that the side-effect effect is present for knowledge ascriptions just as it is for ascriptions of intentional action (Beebe and Buckwalter, 2010, p.476).

[186] Participants were undergraduate students at the University of Buffalo, recruited to participate in the study in a classroom setting.
[187] Participants answered using a seven-point scale anchored from −3 (The chairman didn't know) to 3 (The chairman knew).

Specifically, the folk ascribe knowledge differently in response to changes in the moral valence of the side-effect. Put another way, morally bad side-effects are more likely than morally neutral or good side-effects to yield ascriptions of knowledge.

Thus we have empirical evidence suggesting that moral information plays a central role in how the folk attribute knowledge to an agent. To be clear, this information is epistemically irrelevant and I suggest that epistemologists would typically balk at the suggestion that this information ought to play a role in how we assign knowledge. Not so for the folk.

It is important to emphasize the relevance of this study. Beebe and Buckwalter have discovered that the side-effect effect is present not just in intentional action attributions but also in attributions of knowledge and so once again moral information is playing a significant role in how the folk utilize the concept of knowledge. Importantly these observations help make sense of some other empirical results. Nichols and Ulatowski (2007) solicited open ended responses to why people responded as they did to the Chairman vignette, trying to see how the folk understand their own use of the concept.

As it turns out, Beebe and Buckwalter's study actually helps to explain the pattern that appears from Nichols and Ulatowski's study. In particular, open ended responses to the harm vignette often included explicit references to the knowledge of the Chairman, but the same was not true when responding to the help vignette. In effect, when responding to the Chairman vignette in an open ended format, participants would cite the knowledge of the Chairman in the harm case to explain their ascriptions, but refrain from citing the same knowledge in the help case. These open ended responses mirror the results that Beebe and Buckwalter found, falling in line with the asymmetry in knowledge ascriptions they identified.

These confirming results paint a clear picture: the folk concept of knowledge is ascribed in a way that is significantly shaped by moral information. Again, we have empirical evidence to suggest that our folk psychological competence with respect to the concept of knowledge is shaped in fundamental ways by moral considerations as opposed to quasi-scientific explanatory and predictive considerations. Some critics might wonder how central or important the concept of knowledge is to folk psychology, but when these results are combined with the observation that intentional action and causal responsibility operate similarly, a general pattern begins to emerge. While these empirical results regarding each concept on its own might not be that significant, together they begin to tell a consistent story about the way our folk psychological practice functions. Moreover, and as discussed above, the relationship between knowledge and belief is a close one and so we might expect the asymmetry above to influence folk attributions of belief, an undeniably central concept in the history of philosophical discussions of folk psychology. In fact, there is no need to speculate as this is an empirical question and it has been answered.

5.3.4 – Beliefs and Desires

If knowledge ascriptions are attributed asymmetrically in response to differences in moral valence, we might expect that ascriptions of belief are made in a similar way. If one knows something and so also believes it, we might expect that if moral information influences how we ascribe knowledge, that it might also influence how we ascribe beliefs.[188] While the reasoning behind this speculation is justifiable, once again it's better to simply investigate the matter than think about it.

[188] Of course this need not be the case. For example, moral information might influence the justification process that underlies our ascriptions of knowledge but that need not apply in the case of ascribing belief. This is just to suggest that we might expect the influence of some factor on knowledge ascriptions to ultimately influence ascriptions of belief which are often intimately linked to ascriptions of knowledge.

Results from a study conducted by Tannenbaum, Ditto and Pizarro (2009) having bearing on this matter.[189] Using the familiar Chairman vignette, they explored how the folk[190] would attribute belief to the Chairman in both the harm and help variants. Specifically, they used the following probes to solicit responses[191] regarding the beliefs of the Chairman:

- Did the chairman believe his actions would harm the environment?
- Did the chairman believe his actions would help the environment?

As we would expect, Tannenbaum et al.'s study confirmed the expectation above. Participants' willingness to attribute belief to the Chairman was significantly higher in the morally bad scenario when compared to the morally good scenario[192] (Tannenbaum et al., 2009, p.6). In other words, willingness to ascribe belief to an actor is responsive to moral information regarding a side-effect of their action. These results suggest that yet another folk psychological concept shows sensitivity to the moral valence of a side-effect of the agent's action when used in practice. Importantly, this is not just any concept, but one of the most central concepts to folk psychology, or at least traditional characterizations thereof.

While beliefs themselves are of crucial importance to theorizing about folk psychology and are often investigated in isolation, more often than not in folk psychological accounts of behaviour beliefs operate in conjunction with desires. Recall the familiar explanation of my walking to the fridge to get a beer that involves both my belief (there is a beer in the fridge) and

[189] The primary purpose of Tannenbaum et al.'s study was to explore the effect of "protected values" on the side-effect effect. More specifically, they set out to examine whether distinction in the value placed on environmental welfare would influence judgments of the Chairman's actions. As such, when reporting the specific results of each probe, they do so by breaking out the result into the two comparison groups (i.e. those with a protected value regarding the environment and those without). That said, while they do not report the specific findings for the entire sample, they do discuss the results and comparisons between the harm and help variants at this level.
[190] Participants in the study were recruited from undergraduate courses in psychology and economics at the University of California, Irvine and invited to complete the survey online.
[191] Participants responded using a 7 point Likert scale anchored from 1 "Not at all" to 7 "Completely"
[192] Mean scores for the protected value and non-protected value groups were 6.35 & 5.69, respectively, in the harm variant and 4.21 & 4.43, respectively, in the help variant. Tannenbaum et al. report conducting a 2x2 (protected value status x valence of condition) ANOVA which revealed that a main effect for condition did exert a strong influence on the attribution of belief ($p < 0.0001$) (Tannenbaum et al., 2009, pp.6-7).

my desire (that I have a beer). Given that beliefs and desires often work in conjunction to produce a behaviour it is worth knowing whether moral information only influences our belief attributions, or whether it influences our attributions of desires as well.

Again, results from the study conducted by Tannenbaum et al. helps us to answer this question as they included a probe in their study to explore how desire would be attributed to the Chairman. More specifically, they invited participants to respond[193] to the following probes:

- Did the chairman have a desire to harm the environment?
- Did the chairman have a desire to help the environment?

Results suggest that once again, participants were significantly more likely to attribute a desire to harm the environment to the Chairman than a desire to help the environment[194] (Tannenbaum et al., 2009, p.6) although in both cases, willingness to attribute the desire was quite low. That is, generally speaking participants' responses to these probes fell below the midpoint for both the harm and help variants, in contrast to the attributions of belief where ratings were above the midpoint for both variants. While low willingness to attribute a desire to the Chairman in this case may be an interesting result in itself, the fact that the asymmetry in the willingness of the participants to attribute a desire to the Chairman persists in spite of this, is the more telling result. Even when we are not likely to attribute a mental state to an agent, this willingness is significantly influenced by the moral standing of the agent's action.

I do not think the importance of these results can be understated. Given the central role that the concepts of belief and desire have played in theorizing about folk psychology, these

[193] The same 7 point Likert scale that was used to assess belief was used for this probe as well.
[194] Mean scores for the protected value and non-protected value groups were 3.68 & 2.70, respectively, in the harm variant and 1.63 & 1.57, respectively, in the help variant. The same 2x2 ANOVA was conducted demonstrating a significant main effect ($p<0.0001$) with respect to attributions of desire (Tannenbaum et al., 2009, pp.6-7).

results suggest that at its core, our folk psychological competence is shaped by moral considerations and the pool of concepts that are subject to this influence continues to grow.[195]

5.3.5 – Other Folk Psychological Concepts

Early in the development of this research program some worried that the results Knobe collected with respect to intentional action were a mere quirk of that concept. In fact, McCann (2005) utilized the Chairman harm vignette to investigate whether participants would say the Chairman intentionally harmed the environment, intended to harm the environment and had the intention to harm the environment. In effect, he wanted to test a variety of concepts related to intentional action to see whether willingness to attribute each of the concepts was stable or whether there was variation.

Interestingly, he found that willingness to attribute each of these concepts to the Chairman varied.[196] That is, while a majority of participants were willing to say that the Chairman intentionally harmed the environment, only a modest proportion of respondents were willing to say that he intended to harm the environment, and only a small proportion say that he had the intention to harm the environment (McCann, 2005, p.740). These results show quite clearly that willingness to ascribe each variant of intentionality to the Chairman varies. While this is important in its own right, showing that there may be important differences in how these concepts are used in folk psychology, the results are also limited as McCann did not include the

[195] One might worry that the Chairman vignette is somehow strange or peculiar in a way that leads to the asymmetrical effect in all of these concepts, whereas other (perhaps real world) scenarios might not produce the same type of response. This is a fair critique as many of the results discussed all focus on the particular structure of the Chairman vignette. While I'm willing to concede that further empirical research using different vignettes would support the general conclusion, it is telling that this effect appears at all. Combine this with the fact that causal responsibility and a number of concepts in the subsection 5.3.5 - Other Folk Psychological Concepts below were investigated using a number of different vignettes, I think there is good reason to suspect that the effect is widespread for multiple concepts.

[196] The experiment took place with a large introduction to philosophy class at Texas A&M University.

help vignette in his study and so he was not able to make any claims about the asymmetrical effect Knobe discovered.

That said, Knobe (2004) did test for the asymmetrical effect with respect to the concept of intention.[197] While his results confirmed McCann's observation that the folk are generally unwilling to say that the Chairman had the intention to harm/help the environment, he discovered that willingness to attribute the concept to the Chairman varied significantly depending on the moral valence of the side-effect (Knobe, 2004, p.185). That is, the same asymmetrical effect was observed. In particular, while willingness to say that the Chairman had the intention to harm the environment is quite low, willingness to say that the Chairman had the intention to help the environment is even lower. In fact, with respect to the question "Was it the chairman's intention to help the environment?" no participant reported a positive answer (Knobe, 2004, p.185). These results are similar to those explored above with respect to the concept of desire. While willingness to attribute the concept may be low, the asymmetrical effect is still very much present. Ultimately, this is the important result since this is the result that demonstrates that moral information is influencing the attribution process even if overall willingness to attribute is low.[198] Moreover, Tannenbaum et al. (2009) confirmed this finding of asymmetrical attributions in their experiment using the probe "Did the chairman have an intention to harm/help the environment?" (p.6).[199]

[197] Knobe recruited participants by approaching people in a Manhattan park.

[198] It is worth emphasizing that the result we are interested in is not whether or not study participants are willing to ascribe the concept, but rather whether their ascriptions vary between the morally good and morally bad vignettes. Moreover, it does not matter where on the scale the ratings fall (i.e. whether they are both above or below the midpoint, or whether they straddle the midpoint), the only result that matters is that there is a significant difference in the willingness to attribute the concept that tracks the differences in the moral valence of the side-effect.

[199] Mean scores for the protected value and non-protected value groups were 3.95 & 3.22, respectively, in the harm variant and 1.69 & 1.75, respectively, in the help variant. The same 2x2 ANOVA was conducted demonstrating a significant main effect ($p<0.0001$) with respect to attributions of desire (Tannenbaum et al., 2009, pp.6-7).

As if this was not enough to confirm the analysis provided, Cushman's (2010) study explored above also tested whether ascriptions of an intention to an agent would vary in morally good and morally bad versions of otherwise identical stories. Recall that Cushman conducted a series of studies with morally good and morally bad variants of three similarly structured scenarios resulting in the death of a single person. Using the probe "To what extent did [agent's name] intend for the one to die?" Cushman found the same asymmetrical pattern with higher ratings being volunteered for the morally bad variant of his scenarios than the morally good across all studies (Cushman, 2010, pp.12, 16, 19, 23).

Tannenbaum et al. (2009) also used the Chairman vignette to investigate how the folk attribute awareness to others. Specifically, they assess agreement with the following two probes:

- Was the chairman aware that his decision would harm the environment?
- Was the chairman aware that his decision would help the environment?

Once again, the results of this study suggest that there exists an asymmetry in the willingness of the lay person to attribute awareness to an agent that is responsive to differences in the moral valence of an outcome of their behaviour.[200] That is, whether the folk are willing to say that someone is aware of the outcome of their action depends in part on whether that outcome can be viewed as morally wrong or not. This is a telling result as there may be a strong relationship between attributions of awareness and attributions of belief, desire, knowing, intending and perhaps other concepts.

Pettit and Knobe (2009) also investigated a host of other concepts to see whether or not the side-effect effect would appear in these instances as well. To begin they tested a number of concepts or expressions that embody a positive attitude. More specifically: deciding, being in

[200] Mean scores for the protected value and non-protected value groups were 6.78 & 6.56, respectively, in the harm variant and 5.67 & 5.96, respectively, in the help variant. The same 2x2 ANOVA was conducted demonstrating a significant main effect ($p<0.0001$) with respect to attributions of desire (Tannenbaum et al., 2009, pp.6-7).

favour of, and advocating for. They also tested being opposed to, to see whether the same results would arise for a negative attitude. These may appear to be strange concepts to test and Pettit and Knobe (2009) admit that "[g]iven that the effect had emerged for so many other folk-psychological concepts, we predicted that we would be able to find it even if we simply selected arbitrary expressions that in some way indicated that an agent was holding or displaying a positive attitude toward a given outcome."[201] (p.592)

I disagree with Pettit and Knobe that these concepts or expressions are merely arbitrary, instead, I suspect that having an attitude towards some state of affairs or making a decision between options of action, are all important parts of our folk psychological framework and may play a central role in how other concepts are ascribed or attributed. For example, knowing that someone is in favour of an outcome, event, or state of affairs may lead to attributions of desire. Regardless of the precise role these concepts play in folk psychology, if they have some role to play then empirical results concerning how they are used are of value to this discussion.

To test[202] whether the folk psychological concept of deciding demonstrated the same side-effect effect, Pettit and Knobe reused the Chairman vignette with the following probes:[203]

- The chairman decided to harm the environment
- The chairman decided to help the environment.

Not surprisingly, the same asymmetry was observed. Participants were significantly more likely to say that the Chairman decided to harm the environment than they were to say that the Chairman decided to help the environment (Pettit and Knobe, 2009, p.592).[204]

[201] This quote is in reference to all the tested items excluding 'being opposed to', but the point remains even when considering the negative attitude which still has the same nature of a pro-attitude, just simply in the different direction.
[202] Participants were recruited from an undergraduate philosophy class at the University of North Carolina – Chapel Hill.
[203] Standard 7 point Likert scales were used with anchor points at 1 (disagree) and 7 (agree).
[204] Pettit & Knobe report the mean rating for the help variant as 2.7 and for the harm variant as 4.6 with the difference being statistically significant, $t(35) = 2.4$, $p < .05$ (Pettit & Knobe, 2009, p.592).

To test[205] how the folk use the concepts advocated and in favour of, Pettit and Knobe developed a new scenario. As with the Chairman vignette, two versions were developed, one with a negative moral valence and one with a positive moral valence. The harm variant is as follows:

> The management of a popular coffee franchise held a meeting to discuss a new procedure for preparing and serving coffee.
>
> The assistant manager spoke forcefully in favor of adopting the new procedure, saying:
>
> I know that this new procedure will mean more work for the employees, which will make them very unhappy. But that is not what we should be concerned about. The new procedure will increase profits, and that should be our goal. (Pettit and Knobe, 2009, p.592)

The help variant was nearly identical, except that in this case the assistant manager argues in favour of a policy that would entail *less* work. As such, the vignette was identical except for the last section:

> I know that this new procedure will mean less work for the employees, which will make them very happy. But that is not what we should be concerned about. The new procedure will increase profits, and that should be our goal. (Pettit and Knobe, 2009, p.592)

Participants were randomly assigned to the harm or help condition and then randomly assigned to a particular probe category, that is, either advocated or in favour of. The following probes[206] were used to explore willingness to attribute each concept to the Assistant Manager:

- The assistant manager advocated [was in favour of] making the employees do more work.
- The assistant manager advocated [was in favour of] making the employees do less work.

[205] Participants were once again recruited from an undergraduate philosophy class at the University of North Carolina – Chapel Hill.
[206] Standard 7 point Likert scales were used with anchor points at 1 (disagree) and 7 (agree).

Once again, agreement with the probe in the harm (i.e. more work) variant were significantly higher than in the help (i.e. less work) variants.[207]

Lastly, to assess how the folk attribute a negative mental state, being opposed, they developed another scenario with two variants. The harm variant is:

> The CEO of a company was talking with his assistant. The assistant said: 'We have conducted an in-depth study of the company's proposed new policy. Our study shows that the new policy would decrease profits for the company and that it would also harm the environment.'
>
> The CEO said: 'Look, I don't really care about what happens to the environment. What I care about is making sure that our profits don't decrease. So, with that in mind, let's definitely not implement that new policy.' (Pettit and Knobe, 2009, p.600)

The help variant was identical except the word 'harm' in the vignette was changed to 'help'. To test the attribution of the concept of being opposed to something, the following two probes were presented to participants:[208]

- The CEO was opposed to harming the environment.
- The CEO was opposed to helping the environment.

Yet again, the same asymmetrical result was observed but this time in reverse. That is, participants were significantly less likely to say that the CEO was opposed to harming the environment than they were to saying the CEO was opposed to helping the environment.[209]

In each of the above cases Pettit and Knobe found significant differences for the harm and help variants. That is, willingness to attribute each of the concepts or expressions to the Chairman, the Assistant Manager or the CEO varied depending on whether the participant was

[207] Pettit & Knobe report higher mean ratings for the harm variant (4.1 for advocated and 3.8 for in favour of) than the help variant (2.8 for advocated and 2.6 for in favour of) with a significant main effect, $F(1, 58) = 4.6$, $p < .05$ (Pettit & Knobe, 2009, p.593).

[208] Once again, participants were recruited from an undergraduate philosophy class at the University of North Carolina – Chapel Hill. Standard 7 point Likert scales were used for both probes, with anchor points at 1 (disagree) and 7 (agree).

[209] Pettit & Knobe report a mean rating of 2.3 in the harm variant and a mean rating of 3.4 in the help variant, and report that this difference is statistically significant, $t(54) = 2.0$, $p < .05$. (Pettit & Knobe, 2009, p.600).

responding to the harm or help variant. This evidence led Pettit and Knobe to conclude that the side-effect effect is extremely pervasive and might apply to a wide range of concepts. In particular, they conclude that "the impact of moral judgment is pervasive, playing a role in the application of *every* concept that involves holding or displaying a positive attitude toward an outcome" (Pettit and Knobe, 2009, p.593). Moreover, the fact that the result appears for the negative mental state of being opposed also suggests that perhaps this effect is present for any valenced attitude. Together all of these results tell a very similar story, our folk psychological competence looks to be sensitive to whether or not the behaviour in question is morally good or morally bad.

5.3.6 – A Preliminary Conclusion

Throughout this chapter I have investigated a number of folk psychological concepts to see how they are used by the folk, in practice. We have left the armchair, stopped theorizing about how the folk use concepts, and asked them indirectly through experimental means. Across a wide range of concepts, some of which can be viewed as tertiary in nature or on the fringe of the framework and others which are absolutely central to the framework, the same pattern of results are observed: moral information plays an essential and important role in how folk psychological concepts function in practice.

We can no longer resist this conclusion on the basis that only a concept or two exhibit the pattern. Instead, we have seen the same result appear for a range of concepts spanning much of the folk psychological framework. Moreover, we cannot reject these results on the basis that this pattern only arises for unimportant concepts or concepts that are not central to folk psychology, since the same pattern arises for undeniably core concepts such as belief and desire. Additionally, often the concepts explored relate to one another in important ways confirming and

reconfirming the importance, relevance and sheer systematic nature of this effect. Lastly, the concepts explored here likely inform our attributions of the vast majority of folk psychological concepts. In particular, intention, belief, desire, knowledge, causal responsibility, and awareness are all absolutely essential to the folk psychological practice, would likely inform our attributions of any folk psychological concept, and would likely form the core of nearly any characterization of the framework. It simply is not possible to just wave one's hand and make these results disappear any longer; moral information influences our folk psychological practice and this observation needs to be wrestled with. That said, there is one worry that should be addressed before diving into the implications of these results in the next chapter.

5.4 – Normative vs. Moral Violations

The analysis that results from every single empirical investigation above concludes with the observation that moral considerations are playing an important role in how the folk psychological concept in question is attributed in practice. This observation stems from the asymmetrical effects that are observed in response to each of the vignettes used or the two behaviours described in a vignette. However, there is a very important worry that underlies all of this discussion, namely, whether the outcomes or behaviours in the vignettes can be properly considered immoral. Put another way, we might wonder whether the differences between the two variants is a moral difference or whether they are evaluatively different in some other respect. For example, we might wonder whether harming the environment as a side-effect of some action is really viewed as a moral transgression. Or, in the context of causal responsibility, we might wonder that "perhaps we are not testing whether moral valence plays a role at all since there is a real question whether breaking a company policy is a moral violation." (Roxborough and Cumby, 2009, p.212) It is possible that in each of the cases we've explored thus far, participants

view the outcome or the behaviour as being bad, undesirable or simply objectionable, but this does not necessarily entail that they always view it as being immoral and worthy of moral condemnation.

In Roxborough and Cumby we speculated that while the moral valence of the Professor's behaviour would not be viewed as significantly immoral, we suggested that the folk would recognize his behaviour is at least tending towards immoral. Moreover, we speculated that if the vignette was modified such that it included an obviously immoral action, it would simply amplify the effect already observed. We cannot, however, put this worry to rest with mere speculation. The worry that the relevant information driving the observed asymmetry is not truly moral information is an important one and poses a significant and legitimate challenge to the conclusion I just presented.

If these studies are supposed to demonstrate that moral information plays an essential role in how our folk psychological concepts function in practice, then we better be sure these studies actually involve behaviours or outcomes that worthy of moral assessment. After all, if there is one thing that experimental philosophy has taught us, it is that we can never assume something about the way in which the folk will use a concept. Fortunately, there is some empirical evidence to demonstrate that for most of the vignettes discussed throughout this chapter, the folk do judge the behaviour or outcome as warranting at least some moral condemnation and thus, that in each case we really are dealing with moral judgments.

Let's begin with the Professor Smith vignette as this involves what might be construed as a very minor transgression, namely, the improper taking of a pen. Jill Cumby, Ben Fraser, and I conducted a preliminary study that helps address this issue. Specifically, we conducted a study[210]

[210] The results that follow are from an unpublished study (Roxborough, Cumby, & Fraser (2010)). Participants were recruited from an undergraduate philosophy class at York University.

to test whether or not the transgressions in the Knobe and Fraser vignette and the Roxborough and Cumby vignette would be viewed as a moral violation. We presented participants with one of the two vignettes and asked them to rate their agreement[211] with the claim that "Professor Smith's action was immoral."

The original intention of the study was to investigate whether or not judgments of moral valence were different between the two vignettes. Recall, the Roxborough and Cumby vignette re-introduced statistical atypicality into the picture by describing the Administrative Assistant behaviour as being out of norm in this respect and we wanted to whether the re-introduction of statistical atypicality had any impact on the moral judgments of Professor Smith's behaviour.

As it turned out, Professor Smith's behaviour was judged equally immoral in both vignettes.[212] As such, we collapsed the data providing just one result regarding the perceived immorality of Professor Smith's behaviour.

Our results indicated that participants did view Professor Smith's breaking of an office rule as being worth at least some moral condemnation as the average rating was just above the midpoint of the scale of "somewhat immoral".[213] Some may worry that this result was not conclusive as the behaviour was not judged as being outright immoral because the mean rating was so close to the midpoint and so does not provide enough support to the claim that moral judgments influence causal judgments.[214] This concern has merit so we investigated the results further and found that a slim majority of respondents provided a rating indicating that the

[211] A 7 point Likert Scale anchored at -3 (Not at all), 0 (Somewhat), and +3 (Fully) was used.
[212] Mean ratings of agreement were 0.23 for the Knobe & Fraser vignette and 0.45 for the Roxborough & Cumby vignette.
[213] After collapsing the results, the mean rating of agreement was 0.34 or just above "Somewhat" on the 7 point scale used.
[214] An anonymous journal reviewer proffered this worry when reviewing an early draft of the original paper.

behaviour is immoral.[215] In other words, most of the participants in our study were comfortable enough to make an outright condemnation of the Professor's behaviour, at least to some degree. These results and analysis help to alleviate some of the worry we raised in Roxborough and Cumby.

However, it was also suggested that by asking participants to agree or disagree with the claim "The Professor's behaviour was immoral" we skewed the results.[216] The worry runs as follows. Since the question was phrased as to only include the immoral option, the participants may be responding to the fact that the researchers chose simply to ask this question, suggesting that since the question was asked there might be reason to think that indeed the behaviour is immoral. As such, and with a desire to not appear inconsistent with the researchers, the participants would respond in line with their assumption concerning what we thought. This might explain the apparently weak result we found since, so the speculation goes, participants really wanted to say that the behaviour was not immoral, but did not want to appear out of step with the researchers' assumptions.

It is not clear to me that this line of criticism is entirely fruitful. While the commentator's critique has some merit, it would apply broadly to an array of research not just our particular study. In fact, a great deal of research in the public and private sectors includes phrasings such as this. Sometimes you simply want to assess agreement with a particular position and the best way to get at this is by having the statement itself be clear on a position. Moreover, our common every-day practices are consistent with this phrasing. We often ask ourselves or others "Was that wrong?" or "Was Jane's behaviour yesterday strange?" Speaking in this way is familiar and

[215] 54% of the participants offered a rating of 1 or greater when indicating agreement with the claim "Professor Smith's behaviour is immoral" on a 7 point scale from -3 to +3.
[216] This concern was also raised by an anonymous journal reviewer.

commonplace. We don't always raise the two options that are available to us when talking and so in this way the test statement we used is completely compatible with daily speak. That said, to address the concern empirically we ran a follow-up study revising the test statement.[217]

In this study[218] we used all three permutations of the Pen vignette, including again the Knobe and Fraser and Roxborough and Cumby versions, but also including the Sytsma et al. version. As with the previous study, we asked participants to assess the moral status of Professor Smith's behaviour, but this time asking participants to state how immoral or moral they thought Professor Smith's behaviour was using a 7 point Likert scale with a low anchor point of "Very Immoral", a midpoint of "Neither Immoral nor Moral", and a high point of "Very Moral". As with before, we collapsed ratings for all three vignettes as there were not significant differences between them.[219] Confirming our hypothesis, we found a mean rating slightly below the midpoint demonstrating that even without a potentially leading test statement, respondents tend towards condemning his behaviour.[220] It is worth noting that this slight tendency to condemn the behaviour is consistent with the result we originally observed. Given these results I suspected that if we had presented the respondents with a forced choice between immoral and moral, we would find even higher levels of condemnation. In fact, rebasing the results and removing the midpoint respondents from the analysis, we found that a strong majority of respondents chose to condemn the behaviour.[221] Together, the results of these studies help to demonstrate that at a

[217] Again, this is a part of an unpublished study.
[218] Participants were recruited from an undergraduate philosophy class at York University. Participants were randomly assigned to one of the vignettes.
[219] Mean ratings for the moral status of the Professor's behaviour for the Knobe & Fraser, Roxborough & Cumby and Sytsma et al. vignettes were -0.11, -0.40, -0.34 respectively.
[220] Collapsing across all three vignettes, a mean rating of -0.29 was observed falling just to the immoral side of the scale.
[221] Of those who took a non-neutral position, 74% said that the behaviour was immoral (grouping all responses below the midpoint).

minimum, the pen vignette was on the right track and that the behaviour does elicit a moral evaluation on the part of the participant.[222]

Others have done similar and often more robust analyses of the transgression used in their vignettes and can more clearly answer the objection we are exploring in this section. As we saw above, Cushman (2010) constructed morally good and morally bad variants of acts that are clearly worthy of a moral judgment (for example, either directly engaging in an action that leads to the drowning of a swimmer or allowing the swimmer to drown through an action of omission). To confirm his suspicion regarding the moral status of the variants he asked participants to explicitly assess whether the behaviour was moral or not.[223] In all of the studies he conducted, he confirmed that the morally bad and morally good variants did in fact differ in their moral assessments and that the morally bad case was viewed as significantly more forbidden than the morally good variant. This demonstrates that the results he observed regarding judgments of causal responsibility were in fact related to differences in moral judgments, therefore skirting the concern that was raised regarding the Pen vignette.

Similarly, Tannenbaum et al. (2009) explored the moral valence of the Chairman's behaviour empirically to determine the extent to which his behaviour is actually condemned by the lay person. To test this, they asked participants to assess how moral the Chairman's behaviour was on a 7 point scale anchored at "Completely Immoral" (1) and "Completely Moral" (7). Results clearly indicated that the Chairman's behaviour was viewed as more immoral in the harm variant than in the help variant[224] (Tannenbaum et al., 2009, pp.5, 7). Moreover,

[222] It is worth noting that we did not explore the corollary probe to assess whether the Administrative Assistant's behaviour was also viewed as immoral. Had we done this we would have been able to clearly test whether there is a moral difference between the two actors. All we can conclude here is that Professor Smith's taking of the pen is viewed as being somewhat immoral.
[223] He used a 7 point Likert scale with anchors 1 (Obligatory), 4 (Permissible), 7 (Forbidden).
[224] Mean scores for the protected value and non-protected value groups were 1.78 & 2.07, respectively, in the harm variant and 3.22 & 3.93, respectively, in the help variant. Tannenbaum et al. also report mean ratings collapsing

ratings in the harm variant tended towards the low end or immoral end of the scale, while the help variant yielded ratings much closer to the midpoint of the scale.

Together these results help to construct a reply to the critic who worries that the claim that moral judgments influence how we use and attribute folk psychological concepts is not justified since the scenarios do not properly involve behaviours or outcomes that are worthy of a moral judgment. While the transgressions vary in severity across the scenarios, these results help to demonstrate that the behaviour or outcomes in the scenarios are in fact worthy of a moral judgment. The objection is a good one, however, and any future research aimed at exploring the relationship between folk psychology and moral judgements should always include an assessment of the moral judgments the folk make. However, for now, we can put this objection to rest.

5.5 – Conclusion

Throughout this chapter we have explored a number of empirical studies, many conducted by philosophers engaging in experimental research, all supporting the same general conclusion: moral information influences folk psychological judgments. Put another way, moral judgments and folk psychological judgments are intimately linked, with moral information influencing the way we use folk psychological concepts in practice.

I began with a review of Knobe's original investigation into the concept of intentional action where we discovered that differences in moral valence led to differences in ascriptions of intentional action to otherwise identical behaviours. This asymmetry in the ascriptions of intentional action led Knobe to conclude that moral information informs our attributions of

across the two groups for this assessment. More specifically, the mean rating in the harm variant is 1.99 and in the help variant the mean rating is 3.37, the difference in ratings is statistically significant ($t(298) = 9.52$, $p < .0001$) (Tannenbaum et al., 2009, pp.5, 7).

intentional action. While some early objectors to this research suggested that this asymmetry may be unique to just intentional action (Machery, 2008; Nichols and Ultawoski, 2007) this objection turned out to be mistaken. Instead, what we find is that a multitude of folk psychological concepts show the same sensitivity to moral information that intentional action showed in the original studies conducted by Knobe, including causal responsibility, knowledge, belief, and desire.

In response to the data explored, we might be tempted to provide an analysis of the results that explains away the effect.[225] This move, however, seems misguided given the pervasiveness of the effect and the fact that it shows up for a number of different concepts in effectively the same way. Instead, given the regularity with which this effect is found, it would be prudent to develop one analysis that captures the effect for all of the concepts; moral judgments and folk psychological judgments are tied together in a robust and systematic fashion.

This suggestion follows many others who think that there is some general fact about the relationship between moral judgments and folk psychology that can explain the effects recounted above (e.g. Alicke, 2008; Knobe, 2006; Nadelhoffer, 2006; Nado, 2008). There remains a great deal of debate regarding the precise model that best captures this relationship. For example, above when discussing how the folk attribute causal responsibility Roxborough and Cumby began to develop a model of causal responsibility attributions that spoke to the precise role of all the relevant considerations, including both moral and statistical consideration. I am not interested in this type of debate here; how exactly the various considerations come together to produce the attributions we make is not of particular interest to me in this project. Rather, the mere fact that there is a relationship between moral judgments and our folk psychological practice is already

[225] I will say much more about this option in Chapter Six and why we should not pursue it.

important enough to warrant an investigation into the implications of this relationship. All that is required for this project is that I have convinced you that moral information appears to play a significant role in how each of the concepts explored function, and that this gives us reason to think that the entirety of the folk psychological framework operates in a similar fashion. In effect, we simply need to accept that "there is a perfectly general process whereby moral judgments serve as input to folk psychology" (Pettit and Knobe, 2009, p.587).

This general process, however, has significant implications for the way in which we should understand folk psychology. More specifically, it raises questions regarding the merits of the traditional construal of folk psychology since, as I noted in the introduction, from the traditional standpoint whether someone intended a result, believed their behaviour would cause an event, desired a particular outcome, etc. should not depend on the moral status of their behaviour and yet, this is precisely what the empirical research is telling us. Perhaps then our folk psychological competencies really are being shaped in fundamentally different ways than the traditional account assumes.

Chapter Six: Putting the *Folk* Back in Folk Psychology

6.1 – Introduction

In Chapter Two I argued that folk psychology has traditionally been construed as an explanatory and predictive enterprise. Put another way, the primary function of folk psychology is to facilitate the explanation and prediction of behaviour. Furthermore, I argued that this construal of folk psychology is also committed to the claim that our folk explanations and predictions are quasi-scientific in nature. That is, folk psychology and science share the same goal, namely the explanation and prediction of a natural phenomenon (in this case behaviour), and these explanations and predictions are generally thought to be causal in nature.

More specifically, according to this view when we try to explain behaviour our goal is to cite mental states which are possessed by the person whose behaviour we are trying to explain and that combine in a way to cause the behaviour being explained. The goal is to provide an accurate and testable analysis of the causes that led to the observed behaviour. Similarly, when a prediction of behaviour is provided the predictor is making a claim about mental states that are or will be possessed by a person and the behaviour that will be produced on the basis of those mental states. The goal is to provide an accurate and testable prediction about the way the world will be in the future.

In Chapter Three I argued that folk psychology should be investigated through an empirical lens to ensure that our armchair musings are correct. As we saw in this chapter, philosophers' intuitions about philosophical concepts do not always align with those of the folk and so I argued that since philosophers interested in folk psychology are effectively making empirically testable claims about a real world practice, these claims should be tested. In the context of this project, that means testing whether or not the primary purpose of folk psychology really is to facilitate the quasi-scientific explanation and prediction of behaviour.

In Chapters Four and Five I explored empirical evidence related to this precise claim. More specifically, in Chapter Four I explored empirical investigations to support two claims. First, on the basis of empirical evidence I suggested that perhaps the folk are not as good at explaining and predicting behaviour as they are thought to be. This, I suggested, should give us reason to pause and evaluate the claim that our folk psychology earns its keep because of its remarkable success. Second, I reviewed empirical evidence suggesting that a number of social and cultural considerations significantly shape the way in which our folk psychology is deployed in practice. Again, I noted that this should give us reason to pause and assess whether folk psychology really is a quasi-scientific enterprise or at minimum, whether there are a number of considerations, some of which may not be associated with the goal of making accurate explanations and predictions, that shape and drive our folk psychological practice. In Chapter Five I embarked on a similar endeavour, but this time focusing only on moral considerations and their role in shaping our folk psychological practices. Here we found that information pertaining to the moral valence of a behaviour or an outcome of a behaviour significantly shapes the way in which folk psychological concepts are attributed and used in practice. Again, I suggested that these results should give us reason to pause and assess whether folk psychology really is a quasi-scientific enterprise or at minimum, whether in addition to having the goal of crafting accurate explanations and predictions that goals relating to morality also shape and drive our folk psychological practice.

What remains to be seen is precisely what this empirical evidence means for our understanding of folk psychology, and in particular, the plausibility of the traditional construal explored in Chapter Two. On the face of it, I think this evidence gives us reason to be skeptical

about the claims of the traditional construal, but much work remains to be done to determine the real impact of these empirical investigations. In this chapter I'll begin this undertaking.

First, I'll return to some of the evidence discussed in the previous two chapters to motivate the need for a timeout. The evidence collected demands that we take time to pause and assess the veracity of the traditional construal. I'll argue that, on the surface, our failures in folk psychology and the presence of social, cultural, and moral considerations in our folk psychological practices are inconsistent with the traditional construal in a way that warrants our attention. At a minimum, this inconsistency needs to be addressed and we must provide an analysis of this empirical evidence that either reconciles it with the traditional construal or that embraces a new landscape. I'll review four types of responses to this empirical evidence, two that attempt to reconcile this evidence with the traditional construal of folk psychology and two that reject this construal and begin to sketch out a new landscape. Ultimately I will argue that the attempts to reconcile this empirical evidence with the traditional construal explored in this chapter are the least plausible options moving forward, and that while much work remains to be done to understand our folk psychological practices in a new landscape, this is the right way to move the discussion forward.

6.2 – Press Pause

I've said that we should pause to stop and consider whether folk psychology really earns its keep because of its remarkable success and pause to stop and consider whether folk psychology really is a quasi-scientific enterprise or at minimum, whether there are a number of considerations in addition to making accurate explanations and predictions that shape and drive our folk psychological practice. But why should we pause?

6.2.1 – Why Failure Matters

Does it really matter if our folk psychological explanations and predictions fail to get things right, if they fail to accurately capture the causes of the behaviour trying to be explained or fail to accurately predict a future state of affairs? Yes and no.

If the purpose of folk psychology is to facilitate the production of quasi-scientific explanations and predictions then we should expect that by and large this framework regularly produces successful explanations and predictions. Moreover, if one is going to claim that the framework is so successful[226] that it disappears (Fodor, 1987, p.3) or that the framework's remarkable predictive utility explains its adaptive utility,[227] then we should expect that by and large this framework regularly produces successful explanations and predictions of behaviour.

In Chapter Four, I identified some instances where our folk psychology fails to facilitate the provision of successful explanations and predictions. Of course, it is a further question of whether or not the framework is, by and large, a successful one. While it is difficult to use a few isolated cases of failures to make a generalization regarding the entire practice, I think there is some reason to be concerned about the presumed success of folk psychology based on the evidence explored in Chapter Four. More specifically, the effect of the psychological biases that affect how we use and deploy information and the social and cultural considerations that directly shape our folk psychological practice may have a widespread effect on the success of the framework as used in practice by the folk.

---

[226] I should be clear here that Fodor has a very clear conception of success in mind and one that differs from the notion of success that will emerge later in this chapter. It is true that truth and causal accuracy are only one measure of success (e.g. folk psychology could be successful at serving some social function), but this is the form of success that Fodor, and traditionalists more generally, have in mind. Fodor is talking about the predictive accuracy of the folk theory and the way in which the folk theory maps onto the world of behaviour and mental states. In fact, Fodor goes to great lengths to argue that the generalizations of folk psychology are not empty and that these generalizations are the same in kind as those found in the sciences.

[227] See the discussion in Section 2.4 of Chapter Two.

The psychological biases explored in Chapter Four are well-known and pervasive psychological biases that affect the way folk use and deploy information or make decisions in general. As such, it is likely that these biases permeate our folk psychological practice and so may have a widespread impact on the success of the practice. Worse, the social and cultural considerations that shape our folk psychological practices and that were explored in Chapter Four are also likely to lead to some failures in explanation and prediction. If our practice tends to, for example, craft explanations or predictions of behaviour that cast ourselves or others as rational, then it is likely that our practice will lead us astray when the rationality assumption is not appropriate (perhaps quite often).

Consider again the psychologist, acting not as a psychologist but as a lay person in an experiment, and the explanation he provided when solving the rope-tying task.[228] The explanation he provided painted himself as a thoughtful, logical and rational individual that reasoned through the task and discovered the solution. The social demand to craft explanations that cast ourselves as rational may have been a barrier to seeing the true cause of the solution or may have led the psychologist to genuinely believe the explanation he was providing.[229] The former suggestion is further confirmed by the responses provided by those choosing between identical packages of nylons and the fact that they rejected the placement of the item as a plausible explanation to their behaviour. This explanation, while true, does not accord with the standard that folk explanations of one's own behaviour should, at least in some cases, paint one's behaviour in a way that it appears rational. Worse, if these participants were not explaining their

[228] See Nisbett and Wilson, 1977 or Section 4.3.1 of Chapter Four for a summary of this experiment, the findings, and the psychologist's analysis of his own behaviour.

[229] There is an interesting question here about what would have happened had he put on his "psychologist" hat. While it is not clear that he would have been successful in uncovering the causes of his behaviour in this case, it is clear that the scientific psychologist's job is to do just this. It is unclear whether the folk share in this job.

own behaviour but the behaviour of another person, they may very well have provided an explanation that does not paint the individual as rationally choosing between identical objects. As we saw in Chapter Four, observers do not feel bound to describe actors in the same way they would have described their own behaviour in the same situation. This is problematic from the perspective of the traditionalist, as whether one is an actor or an observer should have no bearing on the explanation provided if the goal is to provide a quasi-scientific explanation that aims to pick out the causes of the behaviour. Whether I am explaining my own or your behaviour has no bearing on what caused the behaviour and yet we are prone to provide different analyses of the causes of behaviour depending on whether it is our own or someone else's behaviour that is being described. And this difference is not merely the product of a difference in knowledge,[230] but it is a product of the difference in the assumptions that underlie the explanations given. If we really were quasi-scientific explainers of behaviour, we should be treating ourselves and others as the same and explaining the behaviour in the same way. The influence of cultural considerations is not benign either and this is evidenced by, for example, the research relating to the fundamental attribution error. Here we saw quite clearly that the way in which one's culture shapes the explanations and predictions offered may lead some cultural groups to be blind to the cause of a behaviour and ultimately fail to provide accurate explanations or predictions of behaviour.

In general, if, as Nisbett and Wilson noticed, social and cultural considerations shape our folk psychological practice, there is the *possibility* that these considerations will not be aligned with the goal of providing accurate explanations and predictions and so our practice might be

[230] There is one potentially meaningful difference that arises when explaining my own behaviour or the behaviour of another. There is likely a difference in knowledge of one's own mental states when compared to one's knowledge of another's mental states. But the difference being highlighted here is not one of merely knowledge, but is a difference in how actors and observers are being characterized.

systematically steered in a different direction. As such, we have reason to be skeptical regarding the claim that our folk psychology really is as successful as it is assumed to be.

This move is akin to one that Churchland (1981) employed long ago, although with a different purpose in mind. In his argument against folk psychology Churchland identified some failures of this framework, in particular, explanatory failures. As Chuchland notes, folk psychology fails to provide any explanatory insight into mental matters other than behaviour. For example, folk psychology does not help us understand mental illness, imagination, and sleep (Churchland, 1981, p.73). Of course, this alone does not show that folk psychology is a poor theory, it just moves us into the realm of *possibility* (Churchland, 1981, p.73). Similarly for our discussion here, a few explanatory and predictive failures even with the possibility of a more widespread effect, does not itself show that folk psychology is a poor framework for explaining and predicting behaviour. But it does raise the possibility and this possibility warrants addressing.[231]

At this point one might worry that I've put too much weight on the claim that folk psychology as used in practice must facilitate the provision of explanations and predictions of behaviour. I am willing to concede that what is more important to the traditional construal is not whether folk psychology *does* facilitate the provision of explanations and predictions of behaviour, but whether the *goal* is to facilitate the provision of quasi-scientific explanations and predictions. Stating that the goal of folk psychology is to provide quasi-scientific explanations and predictions does not commit the traditionalist to the further claim that the framework when used by the folk will always achieve this goal. The idea here, to put it plainly, is to note that

[231] In fact, my argument is arguably stronger than Churchland's. Churchland was concerned primarily with failures of folk psychology to provide explanatory assistance with other mental phenomena; that is, phenomena outside the traditionally circumscribed scope of folk psychology. In contrast, my argument here raises the possibility of failure *within* our folk psychological practices themselves.

sloppy theorizing does not entail that the theory itself is a sloppy one. In this sense, it does not matter that we sometimes fail to provide explanations and predictions of behaviour using our folk psychology. While I will address this idea that the folk's performance is somehow poor but the theoretical analysis of folk psychology is still correct in more detail below, a few words about this move are warranted at this stage.

The traditionalist aligns folk psychology with the scientific practice in a number of loose ways which I think cast a shadow of doubt on the idea that poor performance does not undermine the theoretical commitment that folk psychology is a framework that facilitates the explanation and prediction of behaviour. Inspired again by Churchland, it is surprising that after millennia of use, the practice of folk psychology is still such that it is error prone. If folk psychology is a special purpose tool for facilitating the provision of quasi-scientific explanations and predictions of behaviour, we should be concerned that after millennia of use there are psychological biases and social and cultural considerations that (potentially) systematically steer us away from this goal.[232] We've had ample opportunity to identify and correct for these sources of error, and yet they persist. This worry is further amplified if we take a hard-line traditionalist view of folk psychology and view the folk as operating like proto-scientists. Consider again the child scientist view,[233] where it is claimed that children learning and using folk psychology employ and *revise* their theory just like scientists. Consider further, Sellars' suggestion that our folk psychological practice and scientific practice are on the same continuum (Sellars, 1956/1963, pp.182-183).[234] The move to suggest that the failures identified in Chapter Four do not challenge the claim that

[232] It is true that these biases and social and cultural considerations would influence our scientific theorizing as well, especially early scientific theorizing. But the point of the scientific endeavour is to identify and overcome these biases in order to provide explanation and prediction free from these sort of human influences. It is not at all clear that the same is true for folk psychology.
[233] See Section 2.5 of Chapter Two for more information.
[234] Again, see Section 2.5 of Chapter Two for more information.

the goal of folk psychology is to explain and predict behaviour is less plausible if we take seriously the claim that the folk use folk psychology in a way that is similar to the scientist, insofar as the practice is updated and revised in light of failures to ensure that over time it becomes better. Science is, after all, continually in the business of revising its theories on the basis of explanatory and predictive failures. These failures are taken as evidence that either our method or theory are mistaken and so attempts to refine and improve both are undertaken in order to remedy these failures. In fact, the plausibility of this view is weakened further when we reconsider the fact that some of the evidence explored in Chapter Four suggests that we stick to our folk psychological practice in spite of its failures.[235] To borrow a phrase from Churchland, the lack of refinement to correct for these error inducing considerations is "darkly curious" (Churchland, 1981, p.75) if the goal really is to facilitate the provision of quasi-scientific explanations and predictions of behaviour.

In the alternative, even if we concede that sloppy theorizing is not itself problematic for the traditional construal, there is another sense in which identifying some failures in our practice helps to advance our analysis of folk psychology and critically evaluate the assumptions held by the traditionalist.

To say that we fail necessarily involves judging our performance against some standard. That is, there is a gap between some standard of what it means to be successful and our performance. When we identify that there is a gap between our performance and what it would mean to be successful, two things happen. First, we can begin to understand the mechanisms that

[235] In fact, even identifying failures in this respect may be difficult for the lay person because of folk psychology's success with respect to other standards. These failures are not perceived as failures because ultimately they do not matter and so there is no need to remedy them. But the fact that they do not matter speaks volumes about the actual goal(s) of folk psychology when used and judged by the folk.

are shaping our practice in a way that departs from this standard and second, we can ask ourselves whether this is the right standard to be judging our performance against.

An answer to the first implication has already been provided in Chapter Four, but exploring the second implication provides us with a new way to characterize our performance. By using the language of "error" we have adopted a negative view of our performance. But this negative view is only present because of the particular standard that has been chosen. However, when we look at the mechanisms that lead to our failure, they are often real world considerations. That is, some[236] of the mechanisms that lead us to "fail" just are the social and cultural considerations that have been highlighted throughout Chapter Four. But these social and cultural considerations can themselves set a standard against which our folk psychological performance can be compared. For example, we provide those explanations and predictions that our social and cultural groups can understand and are consistent with their expectations about what constitutes an explanation or prediction of behaviour in these particular circumstances. In other words, the norms of our social and cultural groups. Or we create explanations and predictions of behaviour that conform to a standard of rationality, because rationality is a standard that has social capital or value and so we project it on ourselves or others.

[236] I am more skeptical about whether or not the psychological biases identified in Chapter Four could form the basis of a standard against which we would want to evaluate our folk psychological practice. Unlike the social and cultural considerations that shape our practice and likely have some real world value, it is not clear to me that the psychological biases could similarly have a real world value, although I am open to this possibility. In this sense, they seem to retain their status as a bias even if we evaluate our practice against the social and cultural standards. However, that these biases exists is less of a concern if we adopt a position that rejects the traditionalist assumption regarding the folk psychological practice. That is, if we relegate the quasi-scientific standards of the traditionalist to a secondary position in our analysis of folk psychology, the degree to which these biases have a distortive effect on our practice will be reduced. Put another way, the distortion of these biases is less problematic if we view the practice as satisfying some social, cultural, and moral considerations instead of the quasi-scientific standard of the traditionalist since in this normative environment it is less crucial that we deploy information in a precise and rational manner.

Of course there are a multitude of standards against which the success of our folk psychological practice could be judged. The traditionalist assumes that the right standard is one that is quasi-scientific in nature and that focuses on explanation and prediction. By looking more closely at our failures in folk psychology we can identify that some of the mechanisms underlying this failure might themselves relate to a standard that our practice could reasonably be held against. Of course, there are other standards not yet identified through the empirical evidence explored that we could hold our folk psychological practices to as well. For example, we could assess folk psychological explanations and predictions against a standard of aesthetic quality or how well they tell a particular story and there are potentially many more.[237] A discussion of our apparent failures simply helps to illuminate that we've made an assumption regarding the standard against which we should be judged and helps to identify the mechanisms that actually shape our practice and any corresponding standards that our practices could also be judged against and that would not require us to portray the practice as failing.

Naturally, our performance can be judged against multiple standards simultaneously. For example, I personally tell particularly elaborate stories when I engage in folk psychological discourse. In fact, I have been told that I have a flare for embellishment and exaggeration, sometimes crafting the story in a way that is more interesting at the expense of accuracy. These stories I tell have a communicative value as well as an aesthetic value, but they may also satisfy a number of social standards. For example, they likely often provide an account of behaviour that

[237] An example will help so let's consider a familiar one, my going to the fridge to fetch a beer. This behaviour can be explained in a number of different ways, but offering the explanation "I was thirsty for a beer" will satisfy most (or at least a number of) explanation seekers. However, the behaviour could be explained in a way that fulfills this standard and simultaneously fulfills an aesthetic standard. Consider the following explanation: "My thirst for an amber coloured and slightly hoppy craft beer drove me to the fridge to seek out a beer fitting these delightful characteristics and that would satisfy my craving." This explanation both explains the behaviour in the same way as the first and simpler explanation, but it is also crafted in a way to be far more literary and, for lack of a better word, interesting. The aesthetic experience of the explanation is sufficiently more complex and would satisfy some literary standard much better than the first even if both are equally successful in the quasi-scientific sense.

is rational (especially if I'm describing myself or someone close to me) or that locates the source of the behaviour within the agent to assign them responsibility, and all of this happens within the particular cultural framework that my folk psychological practices are shaped by. But of course, some standards might influence our performance in more significant ways than others (e.g. perhaps a cultural standard is more important than an aesthetic standard) and more importantly, not all standards will necessarily be compatible.

Recognizing the plurality of standards that our folk psychological practice can be held to, allows us to re-examine the success or failure of our folk psychological practices in a new light. In one respect, they do sometimes fail. In another, however, they are extremely successful. If we are providing explanations in a way that conforms to what is expected of us, that conforms to the demands of our culture or society, or that involves us in the discursive community, it is appropriate to say that our folk psychological practices are in fact quite successful even if they do not accurately explain or predict the behaviour. This can be borne out in a number of ways. While we'll look more closely at this below, it is worth identifying one example at this point.

Kristin Andrews (2012) takes this point to heart arguing that explanations generally only need to be provided when something happens that is puzzling and the purpose of the explanation is to resolve this state of puzzlement. In this regard, the success of an explanation is not measured by how accurately it describes the state of affairs in the world, but rather, by how effectively it resolves the explanation seeker's affective state. Importantly, this resolution can be successful without providing an explanation that is quasi-scientific in nature and that accurately captures the true causes of the behaviour. In fact, it could be entirely false and still resolve the explanation seeker's affective state.[238] In this way, folk psychological explanation is absolutely

[238] For more information see Section 6.3.3 below.

successful in one respect, and yet entirely unsuccessful with respect to the traditionalist's account of the practice. As such, our failures have helped us identify and come to grips with the fact that how we assess the claim that folk psychology is eminently successful precisely depends on the standard we are using to define success.[239]

Returning to the question that began this subsection: Does it really matter if our folk psychological explanations and predictions fail to get things right? The answer starts with a yes. Our failures matter insofar as the traditionalist has claimed that our practice is remarkably good. A few instances of failure temper this claim, and the potential for widespread and systematic distortion effects in our practice should lead us to be cautious about the plausibility of this claim. However, these failures may not matter if we can successfully claim that the problem is simply due to a failure on the part of the folk, not a failure in terms of our theoretical analysis of folk psychology. While I gave reason to think that this move is not an ideal one I've left the bulk of this discussion for later and have offered an alternative argument granting this concession but demonstrating that it does not save the traditionalist from the critique. Even if the distinction between a good theory and sloppy theorizing is a viable move, our failures do indirectly help us question the appropriateness of the traditional view by helping us to identify that there are a multitude of standards against which our folk psychological practice can be judged. In particular, by identifying the mechanisms that underlie our failure we've been able to identify other standards against which we can judge our practice and view it as successful. This forces a new question upon us, as we must now ask ourselves to which standard(s) should we be judging the success of our folk psychological discourse? This, however, is a question that will be addressed in much greater detail below (see Section 6.3) as there are a number of options on the table that

[239] For example, Fodor may have been absolutely correct about the success of folk psychology, but entirely mistaken about how we should understand success.

must each be carefully evaluated. The point is simply that our failures do matter and they do warrant that we hit the pause button to stop and evaluate the assumptions of the traditional construal in light of the empirical evidence we've gathered thus far.

6.2.2 – Why Social, Cultural, and Moral Goals Matter

When we engage in a folk psychological discourse, what purpose or goal drives our practice? Put another way, what considerations shape how we engage in this practice? Chapter Two offered a traditional answer to this question, that the goal of folk psychology is to facilitate the quasi-scientific explanation and prediction of behaviour, where the aim of the practice was to accurately capture the causes of behaviour. In both Chapters Four and Five I suggested that the empirical results explored should give us reason to pause and assess whether folk psychology really is a quasi-scientific enterprise, or at a minimum, whether in addition to having the goal of crafting accurate explanations and predictions that our practice is shaped by social, cultural, and moral goals. But why exactly should we care about the suggestion that there are social, cultural, and moral considerations that shape our folk psychological practice and why should this lead us to pause and assess the accuracy of the traditional construal of folk psychology?

I think it's clear from the evidence explored in Chapters Four and Five that social, cultural, and moral considerations are shaping our folk psychological practices and this influence appears to be fairly robust. Why this is important to our understanding of folk psychology and why this should lead us to critically evaluate the traditional construal of folk psychology may not be transparently obvious. Moreover, there is no single and straightforward argument to demonstrate that these empirical findings necessitate that we take a skeptical position with respect to the traditional construal. Instead, there are a handful of related arguments that together justify the claim that in light of these empirical findings, we should pause to assess whether folk

psychology really is a quasi-scientific, and thereby causal, enterprise, or whether at a minimum there are a number of competing goals that shape the practice. It should be noted that what follows is not an argument for the rejection of the traditional view, but rather, an argument in support of taking a skeptical position with respect to the traditional view in order to properly assess what our options are and in this way, continues the argument started in the subsection above.

Above I talked about how our folk psychological failures can help us identify the mechanisms that underlie the way folk psychology is deployed in practice. Here it's important to note the converse, that our social, cultural, and moral considerations may shape our folk psychological practices in such a way that they steer us wrong. More specifically, that these considerations can steer us away from satisfying the goal prescribed by the traditional view. As this relationship was reviewed above in some detail, I will not repeat the bulk of this argument. Instead, I simply wanted to note that these considerations matter to our evaluation of how folk psychology is practiced since they may shape the practice in such a way that we fail at achieving the presumed goal of the practice. Moreover, as noted above, there has been ample opportunity to remedy the influence of these considerations and correct our performance and the lack of improvement in this regard is, once again, "darkly curious". While this alone may not be sufficient to warrant taking a skeptical position with respect to the traditional construal of folk psychology, even momentarily, it does start to build the case. In particular, it starts to demonstrate, at least on the surface, that there are considerations shaping our practices that deter us from achieving the presumed goal of folk psychology. At this point we need to evaluate the role of these considerations in order to better understand the implications of this empirical research. This could be simply describing these considerations as being distortive and thus as

resulting in a performance error on the part of the folk, but this discussion will be had in detail in the next section.

There is another related reason that we should care about this empirical evidence. This is a point that I've stated throughout, but will repeat here to emphasize the significance of this empirical evidence. It is possible that the social, cultural, and moral considerations identified throughout this project are distinct considerations that need not be aligned with the goal of providing quasi-scientific explanations and predictions of behaviour. While it is possible that these considerations can be incorporated into, or be just one element of providing quasi-scientific explanations and predictions, this need not be the case. In Chapter Four I introduced this suggestion through a review of Nisbett and Wilson's seminal work on confabulation and their identification of the social and cultural considerations that shape our reports of mental processes. These rules of engagement that shape our folk psychological practice may include causal considerations, but they may not. As such, insofar as these considerations may be distinct and have instrumental value relating to some goal other than providing quasi-scientific explanations and predictions of behaviour, we again have *prima facie* reason for adopting a skeptical position with respect to the traditional construal of folk psychology. More specifically, we have reason to need to critically evaluate the traditional construal of folk psychology as it may not be telling the whole story.

There is also a sense in which much of the evidence explored in Chapters Four and Five is unexpected. This unexpectedness shows up in a number of different but often related ways. First, these results might strike us philosophers as counter-intuitive in some way. It might, for example, just not strike us as relevant to the question of whether or not someone has a particular mental state, whether or not their behaviour is moral or immoral. It does not strike us as intuitive

that being of a different culture would produce different types of explanations and prediction; an explanation is an explanation and a prediction is a prediction.

A statement of this counter intuitiveness often arises when presenting these empirical results in conference settings or in writing. When detailing these results, those of us who embrace them may often be faced with reactions of surprise or an immediate rebuking of the evidence and an attempt to explain it away. I noted a similar reaction in Chapter Two where I recounted that any suggestion that folk psychology is not a quasi-scientific explanatory and predictive enterprise is often met with shock and a flat out rejection of the claim. The reactions to the results explored in Chapter Four and Five are often the same. This is true in my own experience when presenting the evidence collected, but is perhaps most obviously the case when reviewing the literature on the role of moral information in shaping our folk psychological practice. When originally introduced, these findings immediately found resistance and others tried to explain them away (see for example, Adams & Steadman 2004a/b; Driver 2008a/b). I venture that these reactions and attempts to explain away the evidence are just embodiments of a reaction to a counter-intuitive result; these results do not conform to our commitments or intuitions about how folk psychology should operate and so we attempt to discount and explain away the evidence. This unexpectedness or counter intuitiveness warrants some addressing and warrants taking a temporarily skeptical position with respect to the traditional account of folk psychology in order to more fully explore and assess the incongruity between our understanding of folk psychology and the results of an empirical investigation into the practice.

Continuing along the same lines, but delving a bit further into the story, it would seem that if we really were in the business of explaining and predicting behaviour, the easiest thing to do would simply be to ignore all this information. That is, instead of crafting explanations and

predictions that are sensitive to all of these social, cultural and moral considerations if we really wanted to explain and predict behaviour we would just push all this information aside and get to the heart of the matter; the mental states possessed by the person and the resulting behaviour (see also, Knobe, 2006a, p.21). For example, with respect to the Chairman vignettes, this would mean getting straight to the question of whether or not, for example, the Chairman believes he is going to harm the environment or intends to help the environment. That is, based on the information we have about his mental states and his action, does the Chairman have an attitude of belief with respect to the proposition "this action will harm the environment"? Instead of evaluating the moral valence of his behaviour, *just answer the question*. Instead of trying to craft explanations that absolve a loved one of responsibility for an action, just answer the question at hand, namely, did this person act in such a way as to bring about some undesirable outcome? In effect, our goal should be to just identify whether or not a person had a particular set of mental states or whether there were particular environmental causes of a behaviour, and cite them. This other information and these other considerations are extraneous and simply not relevant to the question in front of us. Furthermore, working within a framework that is sensitive to social demands, cultural demands, and that responds to moral information may ultimately make it more difficult to achieve the goal we set out to achieve and we would be better off if we focused on the question(s) at hand. In fact, this information might distract us from our goal and it would potentially take more time to assess all of these different considerations. As such, we should rightly wonder why these considerations would be playing any role at all and if we'd be better off without them if our goal was as the traditionalist would have us believe.

Moreover, it's not obviously clear that these considerations are compatible with the suggestion that folk psychology is quasi-scientific. Suppose for example that we found a group

of chemists interested in classifying chemical reactions.[240] These chemists would obviously be interested in the elemental composition of each chemical, the ways in which elements interact and bond to one another, the role that catalysts play in chemical reactions and whether any by-products were produced as a result of the chemical reaction. Now suppose that this particular group of chemists was also relying on the economic value, cultural significance, and moral implications of the by-product or result of the reaction when trying to classify the chemical reactions. We would not conclude that these social, cultural, and moral considerations would help the chemists achieve their goal of classifying chemical reactions. Nor would it help them understand, explain or predict the chemical reactions. In fact, historically in those instances where social, cultural, or moral considerations did influence scientific explanations (for example, in the medical sciences), attempts were made to remedy this problem as it was seen as distracting from the primary goal of the scientific endeavour. While these influences may have had a role to play for some time, they were identified and eliminated.[241] In effect, we would not think that the social, cultural, or moral considerations ought to play some fundamental role in our understanding of chemical reactions.[242] Similarly, we might not think that the social, cultural, or moral considerations ought to play some fundamental role in our understanding of behaviour. Instead, as in the case of the chemist, it might be right to view this information as having a

[240] Knobe (2006a, p.6) makes a similar analogy with respect to physics and morality. I have adapted and expanded the story with respect to fictional chemists to help elucidate the analysis and to anchor it to a discussion of a broader range of interests, rather than narrowly focusing on just moral interests.

[241] This is most obviously the case looking back over the history of scientific reasoning. For example, morally objectionable behaviours were often seen as being medically out of norm as well (e.g. pathologizing of sexuality); the moral influenced the medical. But of course, over time the scientific method has been refined to eliminate the influence of these superfluous considerations. It's not clear that the same has been true of folk psychology. Moreover, it's not clear that the folk have any interest in overcoming these considerations, nor is it clear that they should.

[242] This is not to say that there is not important value in the economic, cultural, or moral significance of chemical reactions. It is just to say that this has nothing to do with the chemists' primary goal of coming to understand the reactions themselves.

distortive effect on the process if our goal is to provide explanations and predictions of behaviour.[243] If the best course of action is simply to ignore this information and if this information has no proper role to play in a quasi-scientific endeavour, it should be surprising and troubling to discover that this information is shaping our practice, and so supports the suggestion that we should pause to evaluate the traditional construal of folk psychology in light of this evidence.

Of course, the analogy between science and folk psychology is not necessarily a strong one. While a fairly strong position is available, the more amenable position is that folk psychology is like science in important ways, and that the scientific and folk methods sit somewhere on the same methodological spectrum. Recall Sellars' suggestion that the difference is one of, effectively, rigour. In this way, to compare a chemist integrating social, cultural and moral considerations into their analysis to the folk doing the same may be unfair. While it may be inappropriate for the chemist to make this move, the folk may be able to integrate social, cultural and moral considerations into their analysis while still fundamentally being on the scientific spectrum. That is, while it is not permissible for the chemist to do this and still be considered a scientist, it may be possible for the folk to do this and still be considered quasi-scientists. The suggestion here, however, is merely to question whether in light of all the empirical evidence it is even fair to hold the position that there is no difference in *kind* between folk psychology and science. Given the extent to which social, cultural, and moral considerations shape our folk psychological practice, asking ourselves whether folk psychology really is on the

[243] I recognize that this claim butts up against a significant question. That is, a question about the precise relationship between science and normative considerations. While I think it is commonplace to distinguish the scientific endeavour from a normative endeavour, to examine this question in detail is beyond the scope of this project and for the sake of argument will be put aside.

same spectrum as science is a fair question at this point. This is not to provide an answer to this question, but rather just pose it.

Moreover, even if in the face of this evidence we want to claim that the folk and the scientist are operating within fundamentally the same approach, we can still ask another question, namely, whether the folk are also interested in *different* kinds of issues when engaging in their practice. That is, even if there is some overlap between the scientist and the folk, it is an open question whether there is a pluralism of interests among our folk psychological practice that simply is not being captured by the traditional construal of folk psychology. The empirical evidence explored is suggesting that this could be the case and so the traditional construal of folk psychology might not be telling the whole story, and so warrants review.

Looking more closely at some of the specific empirical results, there is another sense in which some of the patterns that emerge are at odds with the notion that folk psychology is a quasi-scientific explanatory and predictive enterprise. More specifically, when the folk engage in folk psychology, their explanations, predictions and attributions of specific folk psychological concepts exhibit an asymmetry that is not obviously consistent with the traditional construal of folk psychology. This is most obviously the case when looking at the actor-observer effect in folk explanations and when looking at the influence of moral information on our attributions of folk psychological concepts.

Recall the evidence explored in Chapter Four where we saw that actors and observers tend to provide different explanations of the same behaviour. This asymmetry is produced as a result of differences in attention, information, and motivation (Knobe and Malle, 2002, pp.12-13). If folk psychology really is a quasi-scientific explanatory and predictive enterprise, then we should not expect there to be this type of asymmetry. Whether an explainer is the actor being

explained or the observer should have no bearing on the explanation being offered. The right (quasi-scientific) explanation is not dependent on whether or not the person offering the explanation is the actor or an observer, and explanations of identical behaviour should be identical.

Further, consider the TT view of folk psychology explored in Chapter Two. Looking more closely at TT we find that this view is committed to the idea that when we ascribe mental states to ourselves and others we are using the same third person theory. In this sense, we should not expect there to be significant differences in the way that the folk describe the behaviour of themselves versus the behaviour of others since both are based on the same theoretical framework. Our access to our own mental states is mediated through the same theoretical construct that we use to assess the mental states of others, and so there is a sense in which we do not have privileged access to our own mental states when compared to the mental states of others. As such, the fact that there are differences between actors and observers at all may be problematic for this view.[244] Similarly, if ST is right we might not expect there to be differences in the explanations of actors and observers either since the mechanism that produces explanations and predictions is identical whether we are talking about ourselves or others. That is, we are using our own offline action control system and imagining ourselves in the situation of the other in order to explain and predict their behaviour so it is not obviously clear why we

[244] This is a general point of interpretation regarding the TT view. I recognize, for example, that one might argue that while the theory being applied in the case of an actor or an observer is the same, the knowledge that we have to inform our application of the theory is better in the case of ourselves than of others. However, some of the ways that the actor-observer asymmetry manifests may escape this reply. For example, whether or not a behaviour is portrayed a rational may not be a question of knowledge. Further, whether an explanation is crafted that locates the locus of causation within the agent or pushes it onto an environmental factor will not always depend on one's knowledge of the situation being explained. Choosing between "Because it's going to rain today" and "I think it's going to rain today" is not always dependent on our knowledge, but is often dependent on the sorts of social considerations being highlighted in this project. So while I concede that the advocate of the TT view may be able to address some of these asymmetries, I am skeptical that they will be able to explain them all away in all instances and maintain the core notion of the TT view that access to all mental states is mediated through this third-person theory.

should be providing significantly different explanations and predictions of their behaviour since we are effectively describing ourselves.[245] None of this is to say that the differences are not compatible with these views, just to say that it is not obviously the case that there is consistency and additional explanatory work must be done to account for these results. But in a general sense, consistency would demand that identical behaviours get identical explanations and predictions but it looks like the folk psychological practice is not beholden to this standard, a standard we should expect to apply if the goal was to be scientific in approach.

Similarly, the role of moral information in shaping our attributions of folk psychological concepts reveals an asymmetry in our practice that is not obviously consistent with the traditional construal of folk psychology. Consider, for example, the Chairman case. From the perspective of the traditional construal of folk psychology, the help and harm variants are identical. The Chairman's behaviour is the same and his mental states are structurally the same.[246] Given this, we should expect the behaviour to illicit identical attributions of relevant folk psychological concepts, including the mental states of belief and desire. Once again, consistency would demand that given that the relevant factors and the behaviour are identical in both variants, that the attributions of folk psychological concepts should be the same in both cases. Of course, what we

[245] Granted there is more flexibility with the ST view than the TT view since on this account there is a step in the process that allows for there to be less accuracy in attributions of mental states to others than to oneself. However, that there is less accuracy does not necessarily explain the asymmetry in attribution based on the same information. It is the asymmetry that is of interest here, not the accuracy. Because the ST system is run the same way for both self and other, while attributes to others may be less accurate it is not clear why the attribution should be systematically *different* than attributions to oneself. If I understand myself as the locus of causation for a particular action, and I am using that same reasoning process to explain the behaviour of someone else who has engaged in the same action it is not clear how the ST system produces a type of explanation that does not locate causal responsibility for the action within the actor when the explainer is the observer.

[246] The Chairman's concern for the implications of the program is the same and his stated desire is the same. The content of the concern is different, but the attitude is the same in both cases. The relationship between the stated concern and desire is structurally equivalent leading to the same action, and so the natural conclusion to draw is that the lack of concern for the foreseen side-effect combined with the desire to produce profits should yield an analysis of intention that does not capture the foreseen side-effect. It is only once we begin to weigh the moral valence of the lack of concern for the foreseen side-effect that we can parse out these two vignettes as being structurally different.

find is that there appears to be other relevant information that drives the attribution practices in a way that produces an asymmetry between the two variants. This asymmetry needs to be addressed and warrants that we adopt a skeptical position with respect to the traditional view, at least temporarily, to assess our options and properly evaluate how we ought to conceive of folk psychology.

Finally, the results of Chapter Four and Five combine to suggest that our folk psychological practice may have a normative element to it. That is, there is a normative purpose to the practice either in addition to or in replacement of the descriptive practice of providing quasi-scientific explanations and predictions that is assumed by the traditional construal. In this sense, our understanding of folk psychological concepts is very different than our traditional understanding of scientific concepts[247] and in particular, the traditionalist's project. Put another way, how the folk use folk psychology to describe and understand one another fulfills a number of normative goals and the evidence explored suggests that our folk psychological concepts are inherently normative in nature. That is, we use folk psychology to describe and understand behaviour in a way that is different than other physical phenomena. In particular, we describe ourselves and others in a way that is different than non-agential behaviour and this separates folk psychology from the quasi-scientific construal in an important and significant sense.

While this analysis is true for much of the evidence explored and in particular the moral analysis developed in Chapter Five, perhaps it is most relevant when looking at how the folk and the scientist understand causation. Given the evidence explored in Chapter Five, it looks as though the folk psychological notion of causation is significantly different from the physical

[247] It is fairly common-place to understand the scientific method as a descriptive project and for the sake of argument I will take this position as to engage in this debate would involve a treatment that cannot be provided here. Whether science achieves this stated goal and whether there is a proper role for normative considerations in the scientific enterprise are separate questions that cannot be answered here.

notion of causation employed in the sciences. The fact that how the folk understand causal responsibility hinges on the moral valence of the behaviour or associated outcome suggests that the folk use this concept in a way that significantly departs from how we typically understand non-agential causation.[248] Worse, some of the evidence suggests that the moral goal implicit in our folk psychology is so strong that judgements regarding non-agential causal factors can be influenced by our agential causal judgements. Recall that Cushman found that judgments of the causal responsibility of a non-agential cause are influenced by our judgments of causal responsibility for agential causes where that cause was associated with an immoral action.[249] The fact that the folk's understanding of agent-causation demonstrates this normative character should give us reason to wonder whether our folk psychological practice is really analogous to science in the ways that were explored and advanced in Chapter Two; there are two very different notions of what it means to cause an outcome being used here.

But this is true for other concepts as well. From a purely descriptive perspective, whether a person intended a result, believed their behaviour would cause an event, desired a particular outcome, etc. should not depend on the moral status of their behaviour. And yet this is precisely what we see. When we use our folk psychological discourse we are not, it would seem, merely explaining and predicting some physical phenomenon like we do with other scientific theories.

[248] Throughout Chapter Five and here too I have claimed that the folk's understanding of causal responsibility is significantly influenced by their moral judgments. Some might worry that the experiments did not actually investigate folk attributions of causal responsibility, but rather, folk attributions of moral responsibility. That is, while the vignettes and the questions were using the terminology of 'cause' the folk we were making moral, not causal, judgments. I note that in these experiments we've used the word appropriately related to the concept. To suggest that the folk have another concept in mind when using the word, in effect, denies the folk the ability to use words the way they want to. By saying that the folk must be making moral judgments while using the term cause, denies them the privilege of using the word 'cause' in the way they see fit. It, in effect, privileges the interests of the philosopher over the interests of the folk and fails to recognize that folk psychology is a descriptive project where we are effectively trying to capture the practice as is, not as the philosopher thinks it should be. To say that the folk are really using a different concept than the term denotes makes it very difficult to tease apart practice from conceptual commitments, instead of allowing the practice to tell us what the conceptual commitments are. This is an issue that I'll return to when discussing the "performance error" critique below.

[249] See Section 5.3.2 of Chapter Five and Cushman (2010, pp.21-23).

We are instead very much engaged in a normative evaluation of the behaviour, a condemnation of the behaviour, or an evaluation of the behaviour against other social and cultural standards. This normative element of folk psychology distinguishes it from the scientific endeavour in an important way and forces us to evaluate whether we can retain the analysis that the folk and the scientist are operating within the same methodological spectrum. If folk psychology really is a normative enterprise then it may not be appropriate to continue to claim that the folk are acting like proto-scientists when they engage in folk psychology.

Returning then to the question that began this subsection: Why should we care that there are social, cultural, and moral considerations shaping our folk psychological practices? The answer to this question is multifaceted. These considerations appear to shape our folk psychological practices in ways that sometimes lead us to fail to provide quasi-scientific explanations and predictions of behaviour and it may strike us as quite strange that there is any role at all for these considerations if our goal is to provide quasi-scientific explanations and predictions of behaviour. That is, that these considerations are given any weight or consideration in how we explain and predict behaviour is counter-intuitive if our goal is to provide quasi-scientific explanations and predictions of behaviour; it would be easier and frankly more appropriate to just ignore these considerations. More importantly, however, it is unclear whether these considerations are aligned with the quasi-scientific construal of folk psychology or whether they can be. It is real possibility that these considerations may be orthogonal to the quasi-scientific project and so folk psychology does not lie on the scientific spectrum. Alternatively, it is a real possibility that folk psychology is a pluralistic framework which may or may not include the quasi-scientific goals posited by the traditionalist. This is most obviously the case when exploring the normative component that these considerations impart on our folk psychological

practice. While it may be possible to integrate these considerations into an account of folk psychology that is fundamentally a quasi-scientific enterprise, there is enough evidence here to warrant at least asking the question as to whether or not there is a difference in kind between folk psychology and science. Aligning these new findings with the traditional construal is not an obvious task and additional work is required before we can be sure that the traditional construal can adequately account for these findings. The point here is simply to note that the influence of these considerations on our folk psychological practices matters and warrants that we hit the pause button to stop and evaluate the assumptions of the traditional by exploring how to best address the empirical evidence collected.

6.2.3 – Pressing Pause

None of what I have said throughout this section is meant to suggest that we must necessarily reject the traditional analysis on the basis of this empirical evidence. All of this is just to say that, at a minimum, we now have reason to be skeptical of the traditional construal and that we should hit the pause button to assess whether folk psychology really is in the business of facilitating the production of quasi-scientific explanations and predictions of behaviour. The evidence explored suggests that we may not be as good at folk psychology as is assumed and that there are a number of considerations that drive and shape our practice that may be or are, at least on the surface, at odds with the idea that folk psychology is as has been assumed.

The traditionalist has assumed that folk psychology is eminently successful, but we have now cast doubt on the veracity of that claim. The traditionalist casts folk psychology as a specialized tool, but the evidence suggests that it may be a multipurpose tool designed to satisfy a multitude of interests. Worse, it looks as though the folk satisfy a number of normative interests in their folk psychological practices, and yet the traditionalist understands folk

psychology as a descriptive enterprise. Similarly, while the traditionalist also describes folk psychology as a causal framework, the folks' understanding of agent-causation departs from the tradition insofar as it appears to have a normative character to it. In effect, the traditionalist conception of folk psychology is, on the surface at least, incongruous with the empirical evidence that has been recounted throughout this project.

This incongruity between the empirical evidence and the traditional construal of folk psychology needs to be addressed and we need to begin to explore what options are available to us now that we've pressed pause. At this point it's not clear where this analysis will take us and through this analysis we must keep an open mind, exploring both attempts to reconcile this empirical evidence with the traditional construal and those analyses that embrace the incongruity, reject the traditional construal, and begin to sketch out a new folk psychological landscape.

Importantly, even being in this position is already a significant departure from the traditional discourse regarding folk psychology. We are now in a position where we must take a skeptical position with respect to the traditional construal of folk psychology and we have shifted the burden of proof onto the traditionalist to defend their previously unquestioned armchair assumptions regarding folk psychology and the purpose of this practice. Even should it turn out that the traditional construal is our best analysis of the folk practice, this will only be after having to give significant consideration to the implications of the empirical evidence collected and presented against this view. This evidence must be addressed and whether we are a traditionalist or a neo-folk psychologist, we must contend with these and future empirical findings, accounting for them in our analysis of the folk practice. With that in mind, let's forge ahead and tackle this incongruity head on to begin to see what will emerge as the dust I've kicked up begins to settle.

6.3 – The Options

So now what? In response to this empirical evidence and the on the surface inconsistency between these results and the traditional construal of folk psychology, what options do we have to address this inconsistency? I think there are at least four general ways to respond to the empirical evidence explored in the previous chapters.[250]

First, in response to the explored evidence and on the surface inconsistency one could maintain that the traditional construal of folk psychology does, in spite of this evidence, accurately reflect the folk psychological practice. More specifically, one could claim that the primary function of folk psychology is to facilitate the quasi-scientific explanation and prediction of behaviour, but simply account for the empirical evidence by noting that people may not be particularly good at using the framework. Put another way, the empirical evidence does nothing more than identify that there may be a number of considerations that distort the practice preventing us from fulfilling the true goal of the practice. In effect, folk psychology works just fine, it is the folk who exhibit a failure in performance. I'll call this the "dig in your heels" view.

Second, a persistent defender of the traditional construal of folk psychology could acknowledge the existence and influence of these social, cultural, and moral considerations, but still maintain that the primary function of folk psychology is to facilitate the quasi-scientific explanation and prediction of behaviour. In other words, while these social, cultural, and moral considerations do shape the folk psychological practice as deployed by the folk, these goals are all secondary to or parasitic on the primary function of folk psychology which is to explain and predict behaviour. In effect, this option acknowledges the evidence, grants that these social,

[250] There could be, of course, other ways of addressing the empirical evidence not explored in this project.

cultural, and moral influences are indicative of a real goal, but then relegates these goals to a secondary and parasitic position. I'll call this the "primary purpose" view.

Third, like the "primary purpose" view, we could acknowledge that when used by the folk, folk psychology has a number of distinct goals, but recognize them as all having their own unique status in the practice. Put another way, we could advocate for a construal of folk psychology that views the practice as being a multi-purpose tool. This view is distinguished from the second view, as there is no claim that the varied purposes are all parasitic on or secondary to a primary purpose, namely, the quasi-scientific explanation and prediction of behaviour. Instead, folk psychology can facilitate each of these distinct goals depending on the circumstance and potentially facilitate a number of them at once and in some combination. In some cases, providing a quasi-scientific explanation may be the dominant purpose. In others, crafting explanations that demonstrate rationality may be the dominant purpose. In others still, we may attribute concepts to others in such a way as to show or demonstrate a condemnation of the behaviour. In some cases, a number of goals will all effectively compete and shape the practice in a particular way. How all this unfolds is an empirical question and one that would require a significant amount of work to build the various models necessary to capture the interactions of each goal. I'll call this the "multi-purpose" view.

Fourth, and last under consideration in this project, we could argue that the social, cultural and moral considerations that have been discovered to shape our folk psychological practice form the primary goals of this practice. Put another way, we can reject the claim that the primary function of folk psychology is to facilitate the quasi-scientific explanation and prediction of behaviour and supplant this goal with a new and normative goal or set of goals. Within this new approach we can, if we so choose, rebuild an account of explanation or prediction. This can

be done, for example, by reframing our understanding of these concepts or by demonstrating how they follow from the normative project of folk psychology. If we pursue this analysis we have, in effect, flipped the traditional construal of folk psychology upside down. Instead of construing the empirical evidence in such a way as to explain it away to save the traditional analysis, this approach construes the empirical evidence as evidence of an effect that primarily shapes our folk psychological practice and requires an entire rethinking of our understanding. The normative goals that have been identified through the empirical investigations are viewed as the foundation of folk psychology and the traditional construal is, at best, secondary. I'll call this the "alternative" view.

The "dig in your heels" view effectively leaves the traditional construal untouched and explains away the empirical evidence explored in the previous chapters. The "primary purpose" view acknowledges this evidence and gives it consideration, but relegates it to a secondary role in our analysis of folk psychology. The "multi-purpose" view embraces the empirical evidence and acknowledges that folk psychology is far more pluralistic than the traditional construal has acknowledged and so departs significantly from this view. The "alternative approach" view is by far the most radical as it fundamentally changes the story, giving the empirical evidence credence and uses this evidence as the starting point for the theoretical conception of folk psychology espoused. Insofar as it acknowledges that folk psychology is an explanatory and predictive enterprise, it does so in a secondary sense.

In what follows I will provide an analysis of each view. I will argue that the first two views are the least tenable of the four and that either of the latter two options prove to be our best bet in response to the empirical evidence explored. Importantly, this is a significant shift in our understanding of folk psychology. These latter two options are significantly different than the

traditional construal explored in Chapter Two and significantly reshape the way in which we think about folk psychology. We have, in effect, left the tradition behind adopting a new folk psychology that more accurately captures the folk psychological practice as deployed by the folk.

6.3.1 – The "Dig In Your Heels" View

The "dig in your heels" view represents an ardent defense of the traditional construal of folk psychology. Inherent in this view is the idea that the empirical evidence catalogued through Chapters Four and Five has no bearing on how we should understand folk psychology. The claim is that folk psychology, properly construed, is still to be understood as a quasi-scientific explanatory and predictive framework and that while the empirical evidence explored may demonstrate that the folk are not achieving this goal, that there may be other considerations shaping and driving their folk psychological behaviour, that there may be a normative element to these other characteristics and the folk's understanding of agent-causation, none of this has bearing on the question of what the purpose of folk psychology is. At the heart of this view, is the idea that the empirical evidence is indicative not of a genuine purpose or effect, but rather, only indicates that there are a number of biases or distortions that affect our folk psychological practice.

As I see it, there are three lines of argument to support this view. First, a proponent of the "dig in your heels" view might argue that despite the evidence collected, folk psychology is on the whole a fairly successful practice. That is, broadly speaking folk psychology does facilitate the quasi-scientific explanation and prediction of behaviour and the results explored are small in number. In effect, being somewhat bad at the practice does not nullify the claim that we are mostly good at it. This suggestion was addressed in briefly in Section 6.2.1 above. More specifically, above I noted that our failures and the role of the social, cultural, and moral

considerations that shape our practice may be more widespread than is obviously the case from the evidence explored; the evidence explored is really just the tip of the iceberg. This is only a partial response, however.

A more comprehensive response begins by recognizing that just because folk psychology might generally or by and large facilitate the production of quasi-scientific explanations and predictions, does not mean that the traditional view is correct. More specifically, regardless of the degree to which, for example, moral information influences our folk psychological practices, it is not sufficient to reply to this evidence by saying that by and large the traditional conception of folk psychology accurately captures the practice. In fact, we can acknowledge that there is a role for social, cultural, and moral considerations in our folk psychology while still addressing the apparent success of the practice as a quasi-scientific explanatory and predictive enterprise without digging in our heels. That is, there is a sense in which both our poor and good performances can be understood within one conception of folk psychology, and within a conception of folk psychology that doesn't require us to explain the evidence away. This is most obvious in the "multi-purpose view" which posits a number of different considerations that could be driving our practice in any given context. Depending on the context, our explanatory performance can be understood alongside, for example, our social performance allowing for good or poor performance on any of these considerations depending on their precise interaction and the specifics of the context in which we are deploying our folk psychology.

A second reason for preserving the traditional view of the folk psychological competence is that this is the best story we have to understand our folk psychological practices. I do not think this line of reasoning is tenable. I noted in Section 6.2.1 above that there are multiple standards to which our folk psychological practices could be assessed against. The "dig your

heels in" view assumes that we should be assessing this practice on the basis of how well folk psychology facilitates the explanation and prediction of behaviour. Of course we need to have reason for preferring this standard or for thinking that this is the standard that our folk psychological practice ought to be judged against. In Chapter Two we explored one potential reason for thinking that folk psychology should be judged against this standard. More specifically, by thinking of folk psychology as a quasi-scientific explanatory and predictive enterprise we can make sense of its adaptive utility and provide an argument for why this practice developed and persisted. However, an argument on the basis of adaptive utility does not support just this one standard of evaluation. Instead, if we acknowledge that folk psychology could have a number of different purposes (and in particular, social, cultural, or moral purposes) then we can consider arguments regarding the adaptive utility of these purposes as helping to explain why this framework was developed and has persisted. This will become more apparent in the sections below, and in particular in Sections 6.3.3 and 6.3.4 where we explore the "multi-purpose" and "alternative" views, but the point here is to note that it is unclear why the traditionalist picture must be the only way we should evaluate our folk psychological performance and competence. It is true that the traditionalist has a ready-made story for why this purpose makes sense, but the fact that there is this type of story is not unique to the traditionalist picture. If we are going to justify why our performance is going to be measured against a particular standard and provide a story to this effect, this story must be developed in such a way as to not also support alternative conceptions of folk psychology.

Finally, and most substantively, another line of reasoning in defense of this view turns on the notion that the empirical evidence collected identifies nothing more than a performance error. That is, while the empirical evidence does identify a real phenomenon, this itself does not

illuminate anything about the folk's conceptual competence, but rather, only identifies how their performance can go wrong. Put another way, sloppy theorizing does not mean that the theory itself is sloppy. In effect, this line of reasoning positions folk psychology as a properly explanatory and predictive enterprise, that gets distorted or fails to be utilized correctly by the folk in their day-to-day practice. The experimental endeavours explored in the previous chapters highlight the folk's performance, but say nothing about the underlying conceptual competence. We are human, after all, and so our ability to deploy an otherwise robust and quasi-scientific framework will be inherently messy. This messiness helps to explain why there are distortions and distractions and ultimately errors in our application of the framework, but this is a practice issue, not a competence issue and says nothing about the underlying framework itself. Importantly, because we can make this distinction between folk performance and folk competence, it is possible for us to claim that the empirical evidence is indicative of the folk's performance, and in this case a performance error, but that the folk's conceptual competence is as is hypothesized by the traditional construal of folk psychology.

However, if we look more closely at the distinction between performance and competence, I think this line of reasoning quickly loses its plausibility as it is not consistent with the purpose and approach of our investigation into folk psychology. To be clear, proponents of this view would have to argue that the folk's underlying conceptual folk psychological competence is consistent with the traditional construal of folk psychology, but that the factors identified through experimentation are altering or distorting this conceptual competence as it gets expressed in practice. There is a breakdown between the conceptual competence and performance due to social, cultural, and moral considerations identified.

For example, Adams and Steadman (2004a, 2004b) raise this worry in terms of pragmatic implicatures. The basic point is that there is a very significant and practical correlation that deeming a behaviour as unintentional entails that we should not blame the agent for his action. Thus, the participants in the Chairman studies may feel obligated to conform to a social practice and respond to the survey questions in a particular way to ensure (consciously or not) that they characterize the Chairman as a proper candidate for blame. In effect, the respondent wants to ensure that the experimenter *sees* the respondent's response as expressing their commitment to the idea that the Chairman is a proper candidate for blame even though, ultimately, they believe that the Chairman acted equally intentionally in both cases. Under this characterization, the respondents are providing responses that do not indicate what they actually think, but rather, just how they think they should perform in a social environment. If this analysis is correct, then the empirical evidence merely identifies an error in performance and says nothing about the underlying competency of the respondent and so does not touch the traditional construal.

This worry has also been raised as an objection to the broader experimental philosophy movement, as explored in detail in Chapter Three. For example, Ludwig (2007) suggests that the variation we see in the folk concepts of knowledge is not indicative of the folk having a different concept than philosophers, but rather merely indicative of a performance error on the part of the folk who in reality, share the same concept with the philosopher. The philosopher is simply better at accessing the true nature of these concepts and deploying them correctly in practice. The parallel to Adams and Steadman should be clear. There is an underlying conceptual competence that both the philosophers and the folk share, but in practice the folk are more susceptible to various factors (e.g. social, cultural environments, etc.) which tend to distort their

performance.[251] This distortion is what is being observed in the experiments, and so the experiments say nothing about the underlying conceptual competence that is presumed to be shared between philosophers and the folk.

I think there are at least five problems with this line of reasoning. First and in general, it is not clear to me why a philosopher's performance is indicative of their conceptual competence, but the same is not true for the folk. Moreover, there is consistency between a philosopher's performance and their presumed conceptual competence, but the inconsistency between the folk's performance and their conceptual competence only arises because their performance differs from the philosopher's performance. If we do not presume that the folk's conceptual competence is the same as the philosophers' then the performance error goes away.

Second, it is not clear why philosophers should have privileged access to the right performance for *folk* psychology. In Chapter Three I argued that philosophers interested in folk psychology are engaged in a descriptive project; a project to understand how folk psychology is used by the folk. We are describing a practice that exists independently of our theoretical commitments, and we are interested in what the practice *is*, not what we think it *ought* to be. In this way, by assuming that the performance of the folk is somehow in error because it does not conform to our theoretical commitments, is disingenuous to the purpose of our investigation. While there will always be an interplay between our theoretical analysis of folk psychology and the empirical evidence collected, we must ensure that our theoretical analysis of folk psychology

[251] At the time of writing I am not aware of any empirical studies regarding philosophers and their performance with respect to folk psychological concepts. Although there is some research to suggest that philosophers do not always conform to the principles espoused in their philosophy. For example, Eric Schwitzgebel (2009) found that ethics books most likely to be borrowed by professors and advanced students of philosophy were more likely to be stolen than other non-ethics books.

is based on observations and we should resist imposing our own conceptual understanding of behaviour onto the folk practice.

Third, it is unclear to me how we could even infer conceptual competence from anything but performance. In effect, experimental philosophers are interested in a sort of Chomskyian analysis of folk psychology; letting the practice speak for itself and inferring the conceptual competence that underlies it. By driving a wedge between competence and performance, I worry that we'll be able to develop analyses of conceptual competence that could depart quite significantly from performance, so long as it serves some theoretical purpose. But if our conceptual competence is inaccessible by empirical investigations regarding precisely those concepts, it is unclear to me what exactly we are holding on to. There comes a point where the reasonable investigator needs to abandon their theoretical commitments in order to develop a theory which is compatible with the empirical evidence in front of him or her and to simply drive a wedge between what we can observe and what we can postulate just to retain our theoretical understanding of the practice we're trying to observe should raise a red flag.

This is not, of course, to say that empirical evidence should always be trusted and that sometimes holding on to our theoretical commitments in the face of contradictory evidence is not appropriate. But as the methodology improves, as the body of evidence grows, and as the evidence becomes more confirmatory, it strikes me that we must reject the notion that the conceptual competence of the folk must be as is traditionally viewed. If we really are in the business of trying to provide a descriptive account of what the folk are doing when they engage in folk psychology, then our goal should be to discover the purpose of this practice not impose a purpose on the practice that requires us to distinguish between performance and competence. In

so doing, the "dig in your heels" view risks departing from the descriptive project that was embarked on in the first place.

Fourth, and all that much worse for the performance error argument, some of the evidence explored in Chapter Five suggests that the influence of moral considerations may really touch our folk psychological competence. Recall that Cushman (2010) found that "the effect of moral judgment on attributions of causation and intention are not limited to the immediate context of the bad outcome, but rather support a variety of productive inferences." (p.24) This observation leads him to reject the claim that the results of the studies we've explored are merely discovering some performance error. To be clear, the evidence here is not about a single concept in a single attributive event. In this limited scope, perhaps the notion of distortion or a performance error could gain some ground. Instead, the empirical evidence collected suggests a much more profound influence that may span both sides of the performance-competence distinction.[252] As such, given the fact that moral information plays a role not only in the specific attribution, but in inferences that are made on this basis, Cushman concludes that moral information must be playing a much deeper role in our folk psychological practices and that in particular, it must be shaping our conceptual competence.

Fifth, even if we grant that the performance-competence distinction can be appropriately made and that the experiments are really just exploring the folk's performance, how do we now make sense of the folk psychological practice and our theoretical understanding of it? Put another way, assuming the folk are acting in error, what should we do about it? It is not like

[252] Alternatively, one could argue that Cushman's results just demonstrate that the folk exhibit multiple performance errors. One when attributing the concept, and one when using the concept in inferential practices. That is, the conceptual competence remains intact, but the folk simply use err in their inferences. In contrast, philosophers would presumably be much better equipped to make these inferences in accordance with the underlying competence that we share with the folk. This effectively pushes the performance error argument back one additional level. While I think of move of this nature is available to the ardent traditionalist, at this point it begins to look disingenuous and stretches what is appropriate theorizing in light of empirical evidence that contradicts one's view.

philosophers can stand back and say to the folk "oh but you should be working to improve your performance to better and more accurately reflect your underlying conceptual competence!" At the end of the day, the folk may simply not be concerned with ensuring their behaviour conforms to the standard that is being imposed on their practice. Instead, the folk might actually resist any admonishment and think that their performance is in fact correct.[253, 254] Alternatively, and as will be explored below, we could adopt a view of folk psychology that embraces the empirical evidence and tries to account for all the evidence as a part of our theoretical understanding of what exactly the folk are up to when they engage in folk psychological discourse. Taking this route allows us view the practice in a way that is consistent with what is being observed and that would not require additional theoretical work to make sense of the disconnect between *our* theory and *their* practice.

The "dig in your heels" view suffers from a number of problems, most of which relate to the distinction between the folk's folk psychological performance and competence. Instead of engaging in the descriptive project that we set out on, the "dig in your heels" view resists the implications of the empirical evidence in order to retain our theoretical commitments. As I noted in Chapter Three, it is an empirical question whether philosophers and the folk share the same understanding of behaviour and so the folk psychological practice needs to be viewed from an

[253] Unfortunately to my knowledge, no experimenters have done "exit interviews" where respondents are offered the traditional attribution and asked whether, upon reflection, they would like to change their answers. Anecdotally, one participant to my own research into causal responsibility once asked me why I was asking such banal questions about pens, administrative assistants and Professors. I replied, explaining the how the traditional construal of folk psychology would have answered the questions associated with one of the Pen vignettes where Professor Smith had been reminded not to take the pens. The participant was stunned that anyone would think the Professor was not causally responsible for the problem. The traditional construal of folk psychology simply did not align with his interpretation of the events before him.

[254] One might object that we could admonish the folk and that they might, upon reflection, change their answers. It is not clear that these reflective answers, in contrast to their immediate and natural answers, are the ones we want if we are interested in the *folk's* folk psychology. It is also not clear that this would be a good thing. If the folk did give up all of these normative interests that shape their practice, they would be losing something important and valuable. I'll discuss this in more detail in Section 6.5.

empirical perspective. In so doing, there comes a point at which it would be disingenuous to steadfastly hold on to the traditional conception of folk psychology if it cannot make sense of the empirical evidence explored thus far. I've argued that this view is not our best bet for understanding the empirical evidence explored and that casting the evidence as minimal in scope or distortive, is not our best bet in our attempt to understand folk psychology as practiced by the folk.

6.3.2 – The "Primary Purpose" View

The "primary purpose" view also represents a defense of the traditional construal of folk psychology. But this defense acknowledges, in a limited sense, that the empirical evidence collected and explored throughout Chapters Four and Five are evidence of a real and genuine effect that needs to be understood in our analysis of folk psychology. In this sense, this view gives more credence to the empirical evidence than the "dig in your heels" view. However, like the "dig in your heels" view, a proponent of the "primary purpose" view ultimately dismisses the empirical evidence arguing that it does not change our understanding of folk psychology as a quasi-scientific explanatory and predictive enterprise.

Inherent in this response to the evidence explored is the idea that while there are, for example, social, cultural, and moral considerations that shape and drive our folk psychological practice, this is all secondary to or parasitic on the primary function of folk psychology; namely, the quasi-scientific explanation and prediction of behaviour. Put another way, while folk psychology might actually be used for a variety of other purposes or in a variety of other kinds of reasoning, these purposes are either secondary to or only possible because of the primary purpose of explanation and prediction. This type of analysis is not uncommon in the history of the debate regarding folk psychology (see for example, Churchland, 1989; Mele & Sverdilk,

1996), but can equally be applied in the present context and in response to the empirical evidence explored.

This analysis of the empirical evidence is, on the surface, quite plausible. If folk psychology does facilitate the quasi-scientific explanation and prediction of behaviour more often than not (as was suggested above), than we might rightly think that this is the *primary* function of folk psychology. This is, effectively, a claim about the scope of our success and its relationship to our understanding of the function of folk psychology. If by and large our folk psychological framework is successful, then we have good reason to think that our framework is being judged against the right standard or, in other words, that our characterization of the purpose of folk psychology is accurate. However, unlike the "dig in your heels" view, this analysis does recognize that these other considerations genuinely reflect other functions for folk psychology. The response to recognizing these functions as genuine, however, is to either claim that these functions are minor in scope or that our ability to fulfill these purposes is parasitic on the product of the primary purpose of folk psychology, that is, the explanation and prediction of behaviour. We might, for example, suggest that these goals can be subsumed under a broader goals of manipulation and control, which is (arguably) consistent with an explanatory and predictive analysis of folk psychology (see for example, Churchland, 1989). In this way, the "primary purpose" view is a superior alternative to the "dig in your heels" view, however, it still fails to provide an adequate account of our folk psychological practices.

The "primary purpose" view holds that the social, cultural and moral considerations shaping our folk psychological practice are either secondary to and/or parasitic on the quasi-scientific and primary function of folk psychology; namely, the explanation and prediction of

behaviour. In what follows I'll treat these as two separate claims and identify the problems with each in turn.

The claim that the social, cultural, and moral considerations identified in Chapters Four and Five are secondary to the primary purpose of explanation and prediction is effectively an empirical claim. More specifically, this is an empirical question regarding the breadth or scope of the effect of these considerations. This argument effectively takes a pessimistic view regarding the scope of influence the social, cultural, and moral considerations have on our folk psychological practice, predicting that these are limited in scope when compared to the traditional purpose. I've already noted that I suspect the empirical evidence collected thus far is just the tip of the iceberg. This is not just mere speculation however. This suspicion is supported by the trend that emerges with respect to the moral analysis advanced in Chapter Five. More specifically, for an undertaking that began in the early 21st century, our understanding of the role of moral considerations in our folk psychological practice has evolved and expanded extremely quickly. Moreover, this effect has been shown for a range of folk psychological concepts including the central concepts of belief and desire. While I am willing to concede that the results found so far appear to come from limited contexts, I am much more optimistic than the "primary purpose" view grants regarding the degree to which the social, cultural, and moral considerations shape and influence our folk psychological practice. If this view was correct, we would expect that a limit to which the contradictory evidence could be collected would be found and at the present moment, it does not look like there is a limit. The evidence advanced in Section 5.3 found that moral information shaped attributions of a very wide range of folk psychological concepts. In fact, we saw evidence that perhaps any folk psychological concept embodying an attitude towards some state of affairs may show this same effect. Ultimately this evidence led to

the conclusion that "the impact of moral judgment is pervasive, playing a role in the application of *every* concept that involves holding or displaying a positive [or negative] attitude toward an outcome" (Pettit and Knobe, 2009, p.593). If true, this suggests that there is no limit to how far this analysis can go and already suggests that the "primary purpose" proponent is mistaken about the breadth of these effects and their position is already significantly undermined. As additional evidence is collected that confirms the effects of these considerations, this will only serve to further erode the "primary purpose" analysis of the empirical evidence.

Relegating these purposes to a secondary nature also complicates our analysis of the empirical evidence. The empirical evidence explored demonstrates that in some contexts, the primary goal of folk psychology *in that context* is not to explain and predict, but rather, to satisfy some social, cultural, or moral goal. I discussed this above in Section 6.2.1 in terms of standards, showing that folk psychology can be held to multiple standards simultaneously, leading to a question regarding which standards might best be applied in a particular context and to what degree. The corollary of this is that there are multiple goals that we can satisfy simultaneously and the precise interaction of these goals will depend on the context. The empirical evidence explored suggests that in at least the experimental contexts explored the particular mix and balance of goals is such that the social, cultural, and moral goals are sometimes primarily driving the practice. That is, the goals that the "primary purpose" view relegates to a secondary role can in fact take a primary position shaping the practice in such a way that it is not plausible to claim that the primary goal in these contexts is the quasi-scientific explanation and prediction of behaviour. This result challenges the idea that explanation and prediction are the primary purpose of folk psychology. At least in some contexts, this is not the case.[255] This problem is

[255] Many of the findings from Chapters Four and Five demonstrate this point, but for illustrative purposes I will highlight a few examples. Recall the study reviewed in Section 4.4.3 of Chapter Four where Americans and East

further compounded by the number of contexts within which this is not the case and by the observation that in some of these contexts we might expect explanation and prediction to be the primary purpose and yet it is not. Worse still, in some cases there is a conflict of goals insofar as the social, cultural and moral goals may actually compete with the goal of explaining and predicting behaviour. So not only do these social, cultural, and moral considerations sometimes occupy the primary purpose of our folk psychological practices, in some instances the goals we are striving to achieve may directly compete with the goal of explaining and predicting behaviour. This shifting or changing of primary purposes depending on context does not accord well with the "primary purpose" view and is better addressed, as we'll see in the next subsection, by the "multi-purpose" view. Instead of trying to reconcile how secondary goals can trump the primary goal or how secondary goals can operate orthogonally to the primary goal, it would be conceptually much cleaner to simply accept that there are a variety of goals that shape the folk psychological practice and recognize that the question of whether goals are primary or secondary is a context specific question; not something that can be stated in general about the practice as a

Asians were asked to write a short paper defending a particular view, and then asked to read the paper of another who they were explicitly told was given the same task. American participants in this study were more likely to say that this other author agreed with the argument put forward than their East Asian counterparts. This result demonstrates that the Western practice of discounting situational or contextual information when ascribing beliefs is very strong and thus, that these considerations can primarily shape our explanatory behaviour. The possibility that cultural considerations can play a primary role in shaping our practices is further evidence by the fact that cultural considerations must be learned over time. More specifically, as we saw in Section 4.4.4 of Chapter Four, children of different cultures tend to provide the same sorts of explanations and then around 8 years of age Western culture begins to significantly shape Western children's folk psychological practices in a way that diverges from Eastern children. Finally, with respect to moral considerations, the studies explored in Section 5.3.1 of Chapter Five regarding intentional action again demonstrate that moral information can primarily shape our folk psychological practices. Not only did we find here that moral information shaped our practices in a way that was inconsistent with the traditional construal, but that these moral interests transcended other social or cultural interests (i.e. gender, education, religious affiliation, English mother tongue, background in moral philosophy had no effect on the finding and the finding appeared for Hindi speakers as well).

whole. This is precisely the sort of analysis that is advanced in the "multi-purpose" view, which will be explored in the subsection below.

With these considerations in mind, it's unclear whether we really ought to understand the influence of social, cultural, and moral considerations as secondary in nature. If the evidence explored is the tip of the iceberg, if the evidence explored suggests that in some contexts these considerations are the primary purpose, and if the evidence suggests that sometimes these considerations are not compatible with the goal of explaining and predicting behaviour, then we have reason to reject this analysis of the empirical evidence as the most appropriate. While this analysis accepts the empirical evidence as being indicative of a genuine effect, in so doing the analysis is fraught with challenges when trying to reconcile the scope and nature of this evidence against the claim that folk psychology is primarily an explanatory and predictive enterprise.

The suggestion that the social, cultural, and moral considerations are parasitic on the primary purpose of explanation and prediction suffers from similar problems. More specifically, insofar as there are cases where the social, cultural, and moral considerations are incompatible with the explanatory and predictive goal, then this analysis is inadequate. First, it is at least in principle possible that our folk psychological practice is shaped by considerations that operate in competition or opposition to the goal of explaining and predicting behaviour. Or more generally and as I've noted throughout this project, there is nothing that says these considerations need to be aligned with the goal of providing quasi-scientific explanations and predictions of behaviour. As such, it's unclear why these goals would *necessarily* be parasitic on the explanatory and predictive goal when they are not necessarily aligned with this goal.

Second, and more importantly, we have seen instances already where there is an incompatibility between the presumed explanatory and predictive goal and the normative goals

embedded in the social, cultural, and moral considerations that shape our practice. The effect the social, cultural, and moral considerations have on our practice often shape it in ways that go against what the traditional analysis would expect and sometimes directly preventing successful explanations and predictions.[256] This conflict between goals is difficult to reconcile if the presumption is that the social, cultural, and moral goals are somehow parasitic on the explanatory and predictive purpose; that the normative goals of folk psychology are somehow parasitic goals of folk psychology even when their target ends conflict. Of course, the extent to which this conflict will have to be reconciled is dependent on the scope of the conflict. However, whether this conflict is extensive or minimal, the existence of the conflict itself creates problems for this analysis since there will be exceptions to the analysis. Moreover, however this conflict is resolved, this resolution must be empirically verifiable given that it ought to accurately describe the way the folk deploy folk psychology in practice.

In fact, we've already seen an attempt (and failure) to do just this. As we saw in Chapter Five, some have attempted to subsume a moral (normative) analysis of attributions of causal

[256] Again, many of the findings from Chapters Four and Five demonstrate this point, but for illustrative purposes I will highlight a few examples. Recall the findings with respect to selecting between identical pairs of nylons (see Sections 4.3.2 of Chapter Four). Here we observed that our explanatory and social interests can act in opposition to one another. In fact, the explanatory interests were rejected in favour of social interests as participants in the study were unwilling to accept the legitimacy of the actual cause of their behaviour because it did not conform the norms of explanation embodied in our social and cultural group. For a predictive example we can look at some of the findings relating to the fundamental attribution error (see in particular Section 4.4.3 of Chapter Four). Recall that we found that Americans and Koreans made different predictions regarding whether an individual would give another person change for bus fare. As we saw, Americans focused on agential causes, in this case personality traits, leading them to make poor predictions of the individual's charitable behaviour as this focus precluded the incorporation of situational information that Koreans were more likely to utilize in their prediction. These findings confirm that cultural interests can lead to false predictions, but false predictions that conform to and satisfy these cultural interests. Finally, moral considerations can directly conflict with explanatory considerations as well. Recall that Cushman (see Section 5.3.2 of Chapter Five) confirmed that not only does moral information effect our attributions of causal responsibility to agents, but to non-agents as well. More specifically, Cushman found that our non-agential attributions of causal responsibility were influenced by the moral valence of an agent's behaviour, where that behaviour and a non-agential factor jointly caused a bad outcome. Of course, from an explanatory perspective whether a behaviour is morally good or bad is of no significance to attributions of causal responsibility and yet this information directly shapes our attributions of causal responsibility to both agents and non-agents in a way that is in opposition to our explanatory goals.

responsibility in descriptive terms, thereby subsuming it under the traditional construal of folk psychology (see in particular Driver (2008a) and the resulting discussion in Section 5.3.2 of Chapter Five). However, as we saw this analysis failed to most accurately explain the patterns of attribution that emerged from the empirical evidence and resulting discussion. That is, the moral analysis more accurately reflects the true nature of the pattern in a way that a purely descriptive analysis could not account for.

Third, even if it is possible to reconcile the social, cultural, and moral considerations with an analysis of folk psychology as an explanatory and predictive enterprise, I suspect that this move will significantly undermine the attempt to save the traditional construal. In particular, while it may be plausible to characterize explanation[257] as serving social, cultural, and moral, and thereby normative, purposes, this characterization will be a significant departure from the traditional conception of explanation held on to until this point. Instead of the quasi-scientific analysis that is presumed to be the case, this reconciled analysis will cast folk psychological explanations as something more akin to social explanation.[258] This suggestion will be explored in more detail in Subsection 6.3.4 below when we look at the "alternative" view. The point is

[257] Here I focus only on explanation because it is less obvious how the same reconciliation can be made for prediction. A prediction either accurately or inaccurately anticipates a future state of affairs in a way that is more rigid than the types of analyses we can provide for explanation.

[258] If we understand scientific explanation in pragmatic terms, there is a sense in which this conception of scientific explanation is consistent with the revised account of explanation proffered here. However, it is not clear to me that the traditional construal of folk psychology is quasi-scientific in this pragmatic sense. More specifically and as discussed in Chapter Two, the traditional construal of folk psychology views our folk psychological practices as quasi-scientific in that they purport to account for the causes of behaviour. They purport to be true and accurate accounts of the behaviour being explained or predicted and in this way are separate from the psychology of the explainer or predictor and the psychology of the consumer of the explainer and predictor. A more robust discussion of the pragmatic theory of scientific explanation, its relationship to the traditional construal of folk psychology and to the new conception of folk psychology being argued for in this project is beyond the scope of the present task. The point here is simply to note that saving the term explanation in such a way that accounts for the social, cultural, and moral purposes that shape our practice would be a significant departure from the traditionalist's view of folk psychology, insofar as it would embody normative functions and its commitment to providing a causal story would be significantly undermined. It is a separate question whether this re-characterization is consistent with a different conception of scientific explanation, namely a pragmatic account of explanation.

simply that in characterizing folk psychological explanation in this fashion, requires that the core commitments of the traditional view be sacrificed in order to save the language of 'explanation' to describe the purpose of the practice. In effect, the difference between folk psychology and science will no longer be one of degree, but of kind even if we retain the language of 'explanation'.

There is one final consideration that applies equally, whether the "primary purpose" view adopts a purely hierarchical analysis of the purposes of folk psychology or the parasitic analysis. In particular, the "primary purpose" view as a plausible analysis of the empirical evidence is supported by the suggestion that folk psychology has adaptive utility because of its explanatory and predictive power. In effect, the "primary purpose" view embeds the explanatory and predictive purpose in a position that helps to explain the adaptive utility of the practice. While this is true, the adaptive utility argument does not uniquely apply to an analysis of folk psychology that posits explanation and prediction as the primary function of the practice. As I briefly noted in Subsection 6.3.1 above, the social, cultural, and moral considerations identified in the empirical evidence could also be characterized as having adaptive utility. This is true whether we position these purposes as being secondary or primary in nature. Regardless of their precise relationship to the goal of explanation and prediction, these goals can have their own adaptive utility that could help to explain the development and persistence of the folk psychological framework. As such, the adaptive utility analysis of folk psychology gives us no reason to prefer the "primary purpose" analysis of the empirical evidence to, for example, the "multi-purpose" view.

The "primary purpose" view initially seems to plausibly account for the empirical evidence while simultaneously recognizing that this empirical evidence is evidence of real and

genuine effects that needs to be understood and addressed by our construal of folk psychology. This separates it, in a meaningful way from the "dig in your heels" view, but also creates problems for the resulting analysis. In particular, this view suffers from problems generally associated with the need to reconcile competing purposes that shape our folk psychological practice or providing an accurate analysis of how these purposes drive the practice in specific contexts. While some of this critique has depended, in part, on an optimistic or hopeful prediction regarding the scope of the empirical evidence, the current state of affairs is at a minimum suggestive that the "primary purpose" view is not our best bet for understanding this empirical evidence and, as such, for understanding folk psychology as practiced by the folk.

6.3.3 – The "Multi-purpose" View

The "multi-purpose" view represents the first significant departure from the traditional construal of folk psychology as it makes no attempt to reconcile the empirical evidence with the traditional construal. Unlike the "dig in your heels" view which tries to explain away the empirical evidence to save the traditional construal and unlike the "primary purpose" view which takes the empirical evidence seriously but relegates the analysis of this evidence to a secondary position in relation to the traditional analysis, the "multi-purpose" view provides an analysis of the empirical evidence that construes folk psychology as being shaped by a number of distinct purposes. More specifically, the empirical evidence is taken as evidence that there are a variety of purposes that drive and shape our folk psychological practice and that these purposes are distinct from the traditional explanatory and predictive analysis that is offered. Adapting from Knobe (2006a), "the claim is that folk psychology should be understood as a tool not only for generating predictions and explanations but also" (p.1) for facilitating a variety of social, cultural and moral

goals.[259] As such, inherent in this view is the suggestion that the traditional construal of folk psychology is mistaken insofar as it provides, at best, a limited analysis of our actual folk psychological practice. In this way, we have moved away from the traditional construal of folk psychology and offered a more comprehensive alternative understanding of our folk psychological practice, one that captures and explains the practice as a whole.

Importantly, this is not a claim that folk psychology is *not ever* explanatory or predictive. Instead, this more comprehensive alternative simply posits that when used by the folk, folk psychology can have a number of distinct purposes and each purpose may have its own unique status in the practice. By this, I mean both distinct from the purpose of explaining or predicting behaviour and distinct from other purposes that drive and shape the practice. As argued in the subsection above, we have reason for thinking that at least some of the considerations that shape our practice are not parasitic on the quasi-scientific explanatory and predictive function of folk psychology. Of course, in some cases it may be that the purpose of folk psychology really is to facilitate the production of quasi-scientific explanations and predictions of behaviour. The "multi-purpose" view simply recognizes that in other cases the purpose of folk psychology may, for example, really be to facilitate some moral interest regarding behaviour, irrespective of our need to explain and predict behaviour. In other cases, a number of explanatory, social, cultural, and moral goals will interact and collectively compete to determine the precise attributions we make and the way in which our folk psychological practice will unfold. Capturing all these cases and in particular, those that fall outside of the scope of the traditional construal, requires that we revisit our understanding of folk psychology and recognize that the traditional construal is

[259] Knobe's (2006a) original claim only made reference to folk psychology as a multi-purpose tool insofar as it also facilitates moral judgments. In this way, I've extended Knobe's original analysis to include social and cultural considerations as well.

mistaken as it is not an accurate portrayal of the entire practice. Put another way, the traditional construal focuses too narrowly on just one aspect of our folk psychological practice. Importantly, not only is the traditional construal ignorant of a wide range of interests that shape our folk psychological practices, but it is ignorant of the interactions between these various interests in any given context.

How exactly all these purposes interact, compete, or collectively or uniquely shape our folk psychological practice is an empirical question and will almost certainly depend on the context and the needs of the situation. In fact, a significant amount of work will be required to refine and articulate the "multi-purpose" view in order to identify all the goals that shape our folk psychological practice and to build models of our folk psychological practice that capture the interaction of these goals in different contexts and for different folk psychological concepts. Some of this work is already underway. Knobe (2006b) has tried to build a model of how the concept of intentional action is attributed and Roxborough and Cumby (2009) attempted to develop a model of the concept of causal responsibility that recognized the role of both explanatory and moral elements in the attribution patterns observed. Notwithstanding the significant work that is still required, the high-level articulation of the "multi-purpose" view offered here already has merit on a number of different fronts making it a plausible alternative to the traditional construal of folk psychology. In fact, some of the arguments in favour of the "multi-purpose" view are those arguments against the "dig in your heels" and "primary purpose" views. With this in mind, I'll be brief where possible.

First and foremost, the spirit of this analysis is consistent with the commitment expressed in Chapter Three regarding the relationship between experimental evidence and folk psychology. There I claimed that philosophers interested in folk psychology are effectively making a

descriptive claim about a real world practice. They were not legislating a particular world-view onto the world, but rather, observing the world and trying to understand a particular phenomenon. As such, I argued that philosophers interested in folk psychology really are interested in observing folk psychology *in situ*, not from the armchair, and allowing the data to speak for itself.

Obviously all data needs to be interpreted, but the claim here is that the data should drive the theoretical analysis and we should not disingenuously impose our theoretical commitments onto the world, forcing a fit where the evidence suggests that the fit is poor. The two attempts to analyze the empirical evidence explored above do just this, either trying to explain the evidence away, relegating it to a secondary (and thereby less important) analysis, or by deriving it from the primary analysis offered. I argued against these moves above effectively noting that the fit between the theoretical commitments and the empirical evidence is poor. The "multi-purpose" view, however, embraces the approach argued for in Chapter Three by allowing the data to drive the analysis of the practice that we construct. More specifically, in Chapters Four and Five I argued that the empirical evidence suggests that there are social, cultural, and moral considerations driving and shaping our practice. The most straightforward way to account for the observed patterns is to suggest that these considerations are in play in our folk psychological practice. The "multi-purpose" view does just this, taking the empirical evidence seriously and as the starting point for the conception of folk psychology that is built on the back of the evidence. In this way, the "multi-purpose" view is more sincerely aligned with the descriptive enterprise that investigators of folk psychology are engaged in than the alternatives explored thus far.

The "multi-purpose" view is also a compromise position, one that recognizes *both* the intuitiveness of the traditional construal and the value of the empirical evidence catalogued. In

this way it is not a complete rejection of the traditional construal and this is actually an argument in its favour. Obviously the folk explain and predict behaviour all the time and folk psychology, quite likely, is the source of this explanatory and predictive prowess. While the empirical evidence explored has highlighted a number of significant exceptions to this rule, in our day-to-day interactions we may very well often be in the business of providing correct explanations and predictions of behaviour. We will, for example, correctly anticipate some future state of affairs on the basis of a prediction or we will articulate an explanation of behaviour that does accurately capture the causes of the behaviour.[260] The fact that the "multi-purpose" view creates a space for this purpose is an argument in its favour since our explanatory and predictive success needs to be accounted for somehow. Just as I argued that the "dig in your heels" view cannot successfully explain away the empirical evidence, it may be inappropriate to explain away the fact that it appears as though we explain and predict behaviour all the time. In this way, the "multi-purpose" view simultaneously accepts the foundational tenets of the traditional construal while taking to heart the results of our empirical investigation. More specifically, that there are limits to our explanatory and predictive success and that there are a variety of considerations that shape our practice. By portraying folk psychology as a multi-purpose tool that supports not only the explanation and prediction of behaviour, but also a variety of social, cultural and moral goals the "multi-purpose" view takes the best of the traditional construal without the limitations of this view, providing a genuine interpretation of the empirical evidence collected in a way that the "dig in your heels" and "primary purpose" views cannot.

[260] This, of course, ignores at least two arguments against the claim that folk psychology could ever accurately capture the causes of behaviour. First, it could simply be that these mental states are not real and so could never be the cause of the behaviour (e.g., Churchland, 1981). Second, while we can construct an analysis of the causes of behaviour, we may never really know whether the specific articulation developed is the actual cause of the behaviour since, after all, a particular event can be described in a number of equally plausible ways (e.g., Davidson, 1963). For my purposes here, I remain agnostic on whether mental state language can ever be used to construct an accurate causal account of behaviour.

The balance struck by this analysis also leads to another point in favour of the view. More specifically, the "multi-purpose" view makes sense of the observation that the *primary* purpose of folk psychology can vary depending on the context and needs of the circumstance. In the discussion of the "primary purpose" view above, I noted that this view was undermined by the observation that in some contexts the primary purpose of folk psychology appeared not to be the traditional explanatory or predictive function. Instead, while in some contexts explanatory and predictive purposes may be primarily driving the practice, the empirical evidence suggests that that in other contexts (namely those explored in Chapters Four and Five), social, cultural, and moral considerations appear to be primarily shaping how the folk engage in folk psychological attribution practices.[261] This shifting of primacy among competing purposes is easily captured within the "multi-purpose" view. This view is not committed to positioning a particular purpose as the primary function of folk psychology, full stop, and instead portrays the practice as a true multi-purpose tool. Similarly, because this view portrays folk psychology as a multi-purpose tool that is driven and shaped by multiple distinct purposes, no challenges arise for this analysis when some of these purposes are in competition with one another. Above I noted that at least some of the purposes that shape our practice may act in conflict with the explanatory and predictive goal insofar as these purposes detract from our ability to explain and predict or have a normative or non-causal quality to them that is not reducible to the explanatory and predictive analysis offered. This, I noted, presents a challenge for an analysis of folk psychology such that these purposes are secondary to or parasitic on the primary goal of explaining and predicting behaviour. This challenge does not arise for the "multi-purpose" view, which is not committed to any particular claim about the relationship between the purposes other than to say

[261] See footnotes 255 and 256 for examples of the primary role of these considerations and when these considerations conflict with an explanatory or predictive goal.

that the purposes can be distinct in such a way that the practice in genuinely a multi-purpose tool. In this way, the "multi-purpose" view can account for both descriptive and normative purposes that shape our practice but capturing them as well aligned, overlapping, or entirely distinct purposes.

A corollary to these arguments is the observation that the "multi-purpose" view provides a much *cleaner* analysis of the practice than is offered by the "dig in your heels" or "primary purpose" alternatives explored above. The multifaceted approach embraces the empirical evidence and we are not required to jump through theoretical hoops in order to explain away the evidence and we are not required to accept that there are quirks or errors in the performance of the folk. Instead of explaining away a significant body of empirical evidence, this analysis provides us with an alternative view that makes sense of the empirical evidence while capturing the intuitive elements of the traditional construal.

Lastly, and as mentioned in the subsection above, the "multi-purpose" view makes sense from an adaptive utility perspective as well. More specifically, each of the distinct purposes that comprise our folk psychological practice could very well have adaptive utility. For example, the social, cultural, and moral functions of folk psychology likely relate to group inclusion criteria and will facilitate cooperation, both of which arguably have adaptive utility.[262] Moreover, because the "multi-purpose" view includes the explanatory and predictive purpose, the traditional adaptive utility argument applies as well. As such, the adaptive utility of folk psychology under this characterization is at least as, if not more, clear than as it is under the traditional construal.

[262] I will say more about this in Subsection 6.3.4 when discussing the "alternative" view.

All of this is not without consequence, however. To reiterate, the proposal here is a significant departure from the traditional construal of folk psychology. More specifically, this analysis recognizes that folk psychology has a broad range of purposes, a range much more broad than the narrow conception offered by the traditional construal. In this way, the traditional construal is portrayed as mistaken as it provides, at best, a limited or restrictive analysis of our folk psychological practice. If we are to accept the "multi-purpose" analysis of the empirical evidence, and I have argued that there is good reason to do so, our story regarding folk psychology has changed.

Under this analysis, folk psychology cannot be understood as *just* an explanatory and predictive enterprise. There are other genuine purposes that drive and shape our practice that we need to account for in our analysis of the folk psychological practice if we are to provide a comprehensive and accurate account of folk psychology *in situ*. Perhaps most importantly, the empirical evidence suggests that there is a normative character to our folk psychological practices that is outside the traditional construal of folk psychology and some of the purposes that shape our practice conflict directly with the goal of providing explanations and predictions of behaviour.

In these ways, we now have a conception of folk psychology that is a significant departure from the traditional construal. Importantly, given the pluralistic nature of this conception of folk psychology we may rightly wonder whether this pluralistic nature is consistent with a view of folk psychology as science-like. More specifically, whether it's even fair to hold the position that folk psychology and science are somehow on the same spectrum. Given the multi-purpose analysis of folk psychology and given that many of these purposes fall outside of or are at odds with the scientific goal, it does not look favourable for this

interpretation. In particular, given the pluralistic nature of the "multi-purpose" view, and the incorporation of normative and non-causal interests into this view, it is fair to say that folk psychology and science may cut nature at different joints. Put another way, that the difference between folk psychology and science is no longer one of *degree*, but of *kind*.

While it is true that a lot of work remains to be done in order to flesh out the "multi-purpose" view in the same ways that the traditional construal has been fleshed out for the past three plus decades, this is not a new or unexpected result when a paradigm shift is underway. Moreover, that the practice of the folk seems to be multifaceted and context specific complicates things greatly and requires careful and nuanced investigations into the practice. But the point here is that our understanding of folk psychology, up to this point, is significantly limited if we are always required to come back to just the one goal of explanation and prediction. Our focus on explanation and prediction does not tell us the whole story and we have work to do if we are going to understand the folk psychological practice as practiced by the folk. But the work that remains to be done is virtuous and productive.

6.3.4 – The "Alternative" View

The "alternative" view represents another significant departure from the traditional construal of folk psychology. This analysis of the empirical evidence also does not make an attempt to reconcile the empirical evidence with the traditional construal. Instead, the empirical evidence is seen as genuine evidence of a number of distinct purposes that shape our folk psychological practice. In this way, the starting point for this analysis and the "multi-purpose" view is the same. However, where the "multi-purpose" view is conservative in its analysis of the empirical evidence, the "alternative" view is more ambitious. While the "multi-purpose" view does not reconcile the empirical evidence with the traditional view, this analysis does create a very clear

and distinct space for explanatory and predictive goals in our understanding of the functions of folk psychology; effectively treating them as equals within a "multi-purpose" account of our practices. The "alternative" view, however, provides a fundamentally different type of analysis than is offered by the traditionalist.

On this construal of folk psychology, the social, cultural, and moral considerations are not just taken as evidence that there are, *at a minimum*, multiple considerations in addition to explanation and prediction that shape our practice. Instead, these considerations are construed as being *foundational* to the practice of folk psychology. In this way we have moved yet another step away from the traditional construal of folk psychology, supplanting this analysis with an alternative. Unlike the other views explored thus far, the claim here is that fundamentally folk psychology should not be understood as a quasi-scientific explanatory and predictive enterprise. Instead, these other goals have a fundamental role in shaping our practice. As such, the traditional view is not rejected just for providing a limited or incomplete analysis, but because it is seen as being fundamentally mistaken about the *primary* function of folk psychology.

Depending on exactly how we cash out this suggestion, the "alternative" view can be looked at as the polar opposite of the "dig in your heels" or "primary purpose" view. This exact relationship depends, in part, on whether and to what extent providing explanations and predictions of behaviour is built back into this new construal.[263] If the quasi-scientific explanation and prediction of behaviour is not given any space within this new account, then we have effectively explained away the intuitive claim that we do seem to engage in these practices. As such, we would be presenting an alternative account that is in direct opposition to the "dig in your heels" view. In contrast, if we create some space within this new construal for an

[263] How explanation and prediction can be built back into the "alternative" view, will be explored in more detail below.

explanatory and predictive function, then we have offered an alternative that is much more like the polar opposite of the "primary purpose" view. Instead of construing the quasi-scientific explanation and prediction of behaviour as the primary purpose and these other considerations as secondary to or parasitic on this primary purpose, the "alternative" view construes the social, cultural, and moral considerations as having a fundamental role to play in shaping our practice and the explanation and prediction of behaviour as being secondary to or parasitic on these purposes.[264]

In the subsection above, I argued that the "multi-purpose" view provides a sensible analysis of the empirical evidence collected; a compromise position. As such, we could stop there; there is no need to adopt an analysis that goes a step further. The goal in our pursuit to understand folk psychology, as I've argued, is not to impose a particular theoretical position on the world, but rather, allow the data to drive the production of our theoretical position. That said, the "alternative" view does offer a genuine alternative to the traditional analysis that construes the function of folk psychology in a way that is consistent with the empirical evidence, just extends this analysis a little further in the direction the evidence is already pointing us. I think this extension warrants consideration.[265] In order to do this, we'll take a closer look at the different ways the "alternative" view can be articulated and how they align with and are well

[264] Granting that explanation and prediction have some role to play in the "alternative" view opens this line of reasoning up to a few criticisms. First, it risks reducing this analysis into the "multi-purpose" view. Second, this account of folk psychology will have to address how explanation and prediction can be derived from the normative (e.g. social, cultural, and moral) primary purpose and yet the converse is not true for the "primary purpose" view, as this was an argument against the "primary purpose" view. I will explore these issues in more detail below.

[265] There are at least two reasons that this extension warrants consideration. First, as I've mentioned, the analysis presented below is consistent with the empirical evidence; it helps to make sense of why we're seeing what we're seeing in practice. But there are also good theoretical reasons for adopting this approach. I will not explore these reasons in great detail in what follows, but the authors of the views about to be explored present a more robust justification of their accounts than is offered here. My goal here is only to explore them enough to demonstrate their compatibility with the empirical evidence; an analysis of the merits of these views independently of their consistency with the empirical evidence is beyond the scope of this project.

positioned to interpret the empirical evidence that has been collected throughout Chapters Four and Five. While there may be many options available to us, there are two views available that I think align nicely with the "alternative" view as I've characterized it.[266]

First, Kristin Andrews' (2012) account of folk psychology warrants consideration as one option within the "alternative" view as she presents a genuine alternative to the traditional construal of folk psychology. While her account of folk psychology is pluralistic in nature and does retain the core idea that folk psychology can have explanatory and predictive purposes,[267] her understanding of these practices departs from the tradition in significant ways, and in ways that accord with the "alternative" view as I have characterized it. In particular, and of most importance for our purposes here, it is Andrews' characterization of folk psychological *explanation*[268] and her rejection of the traditionalist construal of folk psychology that warrants consideration. In her rejection of the traditionalist commitments, Andrews rethinks folk psychological explanation offering a fundamentally different understanding of the purpose of folk psychological explanations.

To begin, Andrews construes folk psychological explanation as involving both "seeking and generating explanatory information." (Andrews, 2012, p.120) In particular, she observes that the folk paradigmatically,[269] or at least typically, seek explanations when a behaviour they've

[266] I should note that I do not believe either of the views explored below were proposed as a way to address the specific empirical evidence I've collected. Moreover, neither view was obviously presented with the intention of being repurposed in the way that I have employed them.
[267] See in particular Chapter 10 and p.184 for an articulation of the principles underlying her pluralistic conception of folk psychology.
[268] It is worth emphasizing for clarity, that I am focusing very narrowly on how Andrews understands folk psychological *explanation*. That is, while Andrews offers a treatment of folk psychology more broadly I am, for the purposes of this section, focused narrowly on just her account of our explanatory practices.
[269] Paradigmatically speaking, we seek out or formulate explanations when we are puzzled by the behaviour observed. Without this state of puzzlement we would, generally, have no reason to engage in explanation seeking or formulating behaviour. In this way the puzzlement is a trigger for our folk psychological practices to be engaged. Of course, we can offer explanations for things that don't puzzle us, it's just that this is less common and so, we might think, less likely to serve as the foundation for the practice.

observed strikes them as anomalous or puzzling. Put another way, folk psychological explanations paradigmatically begin with a state of curiosity or puzzlement in response to an observed behaviour. This affective state then drives the person to engage in explanation seeking behaviour, thereby becoming an explanation seeker (Andrews, 2012, pp.120, 125). Of course an explanation seeker can become bored or distracted and cease their seeking behaviour, but generally they will continue being an explanation seeker until an explanation is found that resolves this affective state. As such, the success of an explanation on this account is not whether it accurately describes the state of affairs in the world that led to the behaviour. Rather, the marker of success of Andrews' account is much more pragmatic in nature. Success is achieved when the affective state is resolved and the explanation seeking behaviour ceases. In this regard, success just simply is the resolution of the explanation seeker's affective state.[270]

Intuitively, this analysis has merit. At the most fundamental level, someone seeking an explanation will continue to seek an explanation until they are provided with an account of the behaviour observed that resolves the state of puzzlement they are feeling. As a corollary to this, the person offering the explanation[271] has just one fundamental goal, namely, resolving the state of puzzlement and bringing the explanation seeking behaviour to an end. Of course, how precisely this is cashed out in each circumstance will vary, but across all circumstances explanation providers will share the same fundamental goal of seeking to resolve the affective

[270] Andrews is also interested in whether the explanation helps to facilitate or maintain group membership and ostracize those who act inappropriately. In my view and for our purposes here, this latter element of the view is important, but of less interest at this stage in the analysis. In my assessment, the more unique proposal offered by Andrews is the analysis of folk psychology as a mechanism for resolving an affective state in the explanation seeker's behaviour and so this will be the focus of much of the discussion that follows. The success of an explanation with respect to defining and maintaining a particular social group is a broader claim that will also apply to the second "alternative" view of folk psychology explored below. Moreover, this element of her view is generally consistent with the adaptive utility argument that I've mentioned throughout this chapter and that I will expand on in more detail below. For these reasons, I've opted to emphasize just the affective state resolution component of this analysis and will just identify the relevance of the group mechanisms in footnotes where warranted or of interest.

[271] It is worth noting that the person offering the explanation and the explanation seeker can be the same person.

state of the explanation seeker. In this way, the alternative analysis being proposed is very much in the spirit of the argument developed in Chapter Three. At the most basic level, this alternative is premised on an observation about how the folk engage in folk psychological discourse. But this analysis of our folk psychological practices also presents a significant departure from the traditional construal.

Importantly, resolving the affective state within the explanation seeker can be achieved without providing an explanation that satisfies the quasi-scientific standard of the traditional construal. First, on this account of folk psychological explanation there is no requirement that the explanation offered and consumed, be true; it just has to be believed to be true. In Andrews' words, the explanation seeker must have the "belief that the explanation is true." (Andrews, 2012, p.124) Once the explanation seeker believes the explanation is true, their seeking behaviour will cease and be replaced with a positive feeling of relief or happiness (Andrews, 2012, p.125). Of course, believing something to be true and something being true may have very little in common. In fact, in some cases, this approach may be directly opposed to the truth, given that the folk may not and need not give much credence to the constraints of rationality or evidence (Andrews, 2012, p.154).

This account also departs from the tradition in its treatment of the causal nature of folk psychological explanations. Andrews' recognizes that while it might be too hasty to think that the folk are prepared to view some actions as not having a cause (Andrews, 2012, p.178), they do construct and consume folk explanations that are not causal in nature; at least not as the traditionalist understands the causal nature of folk psychological explanations.

On Andrews' view, folk psychological explanations can make reference to and be composed of non-causal terms such as dispositions or statistical generalizations and as such,

there is no requirement that folk psychological explanations specify a determinate psychological cause as would be required by the traditionalist conception of folk psychology (Andrews, 2012, pp.125, 143). For example, personality trait explanations can be viewed as causal, but in a counterfactual sense.[272] Andrews rightly observes that at best the conception of causality being deployed here is weaker, or at least distinct from the conception of causality that is embodied in the traditional view. Recall in Chapter Two we saw a clear commitment to the idea that the mental states being cited by our folk psychological practices were the direct cause of the behaviour. There is no similar requirement on this account of folk psychological explanation.

Moreover, Andrews' portrayal of folk psychological explanation also imparts a great deal of relativity to what constitutes an explanation. In particular, what constitutes an explanation will be relative to the explanation seeker. So, for example, what counts as an explanation for me regarding some behaviour, may be different than what constitutes an explanation of the exact same behaviour for you (Andrews, 2012, p.125). Additionally, what constitutes an explanation of a behaviour for me today, may not be a satisfactory explanation of the same behaviour at a different time. In these ways, folk psychological explanations are very fluid. Combine this observation with the observation that truth and causality need not play fundamental roles in shaping folk psychological explanations, and the class of potential explanations for a given behaviour is quite wide and varied. This is not to say, of course, that the class is infinitely wide as there are functional constraints[273] on what constitutes an explanation (Andrews, 2012, p.125), but the class is indeed wide and varied.

[272] Andrews gives an example to help illustrate the point. If we explain a father's reluctance to let his daughter date by pointing to his "old fashioned" or "conservative" personality, we have in effect provided a counterfactual causal story of the behaviour. Were it not for his conservative personality, he would not be reluctant (Andrews, 2012, p.178).

[273] One's social group and the norms that are embedded in this group will also similarly constrain the class of potential explanations, but we'll see this play out in more detail shortly.

At this point there is already a significant disconnect between the tradition and the alternative offered by Andrews. Truth and causality are both of paramount importance to the quasi-scientific perspective and so this disconnect is obvious on just these measures alone.[274] On this view, the core elements of the traditional construal of folk psychology are rejected as being of fundamental importance to the practice of folk psychology as deployed by the folk. Folk explanations of behaviour need not be an accurate account of the causes[275] of behaviour or be truthful; they must just be perceived in such a way by the explanation consumer that their explanation seeking behaviour is resolved. In this way, the construal of folk psychology offered by Andrews significantly departs from the traditional construal, supplanting the traditional purpose with a new and fundamentally different purpose that underlies our explanatory practices.

While I generally find this alternative construal of folk psychological explanation to be plausible in its own right and there are good reasons for adopting this view, I have selected it as an option worthy of exploration in this particular context because of how it accords with the empirical evidence collected in Chapters Four and Five and how it helps to make sense of the conclusions I've been drawing on the basis of this evidence. Put another way, it provides a plausible analysis of the empirical evidence that lets us address this evidence and release the pause button.

First, this view makes sense of our failures to accurately explain behaviour. More specifically, our apparent shortcomings with respect to folk psychological explanation are not viewed as shortcomings on this construal. That is, the fact that we fail to provide explanations of behaviour that are accurate is of no consequence on this view. Unlike the "dig in your heels" or

[274] In fairness, this is just the tip of a very large iceberg, as Andrews' argues against the traditionalist conception of folk psychology on a number of different fronts.
[275] Understood as the causes which directly led to the behaviour.

"primary purpose" view, we do not have to explain away the fact that while our folk psychological practice is (apparently) designed for a specific purpose, that we often fail to achieve this purpose. Instead, the standard of success on this account does not conflict with our observations of practice. As stated above, the success of an explanation is measured simply by whether or not the explanation is consumed by the explanation seeker, believed to be true, and resolved the explanation seeker's state of puzzlement. All of this is possible even when we fail to accurately explain behaviour by explicitly identifying the true causes that led to the behaviour. Importantly, the analysis here is different than what the "multi-purpose" view offers. In particular, the "multi-purpose" view accounts for this issue by allowing different standards to apply in different degrees or in different contexts. On this view, a higher level analysis is offered where the features of what it means to consume an explanation in the most general sense, apply broadly. The standard is the same, believing the explanation to be true, whenever an explanation is consumed.[276]

This also helps to make sense of, for example, the confabulation literature we explored in Chapter Four and in particular, the observation that some of our explanations are just mere fictions. The fact that we provide explanations of behaviour that are in line with implicitly held social or cultural demands, but that are mere fictions, is perfectly consistent with the idea that at rock bottom we are crafting explanations that will be consumed by others. We confabulate because our confabulation will resolve the tension we experience as explanation seekers. The

[276] Of course there could be a plurality of interests that are driving the "frontline" production and consumption of explanatory information. But the point should not be lost, that this analysis can be viewed as a higher level analysis of our entire practice. To perhaps draw this out more clearly, while both the "multi-purpose" view and this particular characterization of the "alternative" view necessitate a change in our research paradigm, Andrews' believes that the research project that her view necessitates will be far too complicated and nuanced to allow for the development of a "flow chart" analysis of folk psychological explanation (Andrews, 2012, p.192). Contrast this to the research project I presented as flowing out of the multi-purpose view where we can try to understand and explore the interaction of all the various standards and interests that are at play in a given context.

explanation offered, while not true and not causally accurate, is offered in such a way that it is designed to be consumed by an explanation seeker and perceived to be true.[277] In fact, the suggestion that the social and cultural rules that shape our folk psychological practices need not be aligned with the scientific enterprise is entirely consistent with Andrews' suggestion that folk psychological explanation need not be aligned with and may be opposed to truth. The fact that folk psychological explanations are crafted in line with social or cultural considerations does not present a challenge for this view. In fact, explanations that conform to social or cultural considerations are primed for being successful if the explanation seeker is a member of those social or cultural groups. A closer look at the cross-cultural variation that was observed in Chapter Four and how this account of explanation makes sense of these practices will be helpful.

More specifically, on this construal of folk psychological explanation it is perfectly reasonable to see variation in what constitutes an explanation of behaviour between people of different cultures. Since there are no *a priori* restrictions on what an explanation must look like in order to be a successful explanation it is possible for different cultural groups to view behaviour differently and consume explanations differently. In fact, culturally different explanations of the same behaviour can both be sound and successful, so long as they are both consumed appropriately by those within the culture.[278] This is also consistent with the explanatory relativity that is inherent within Andrews' view. More specifically, the fact that on

[277] And of course, by conforming to the expectations of one's social group, these explanations help to promote and maintain group inclusion criteria.

[278] Again, this also accords with the suggestion that folk psychology facilitates the creation and maintenance of group inclusion criteria. The influence of cultural considerations in folk psychology is indicative of the importance of producing and consuming explanations that align with one's culture's commitments. If we were to produce explanations that didn't conform to our culture's commitments we would not likely be invited to produce future explanations. The mechanism by which our folk psychological practices constrict behaviour and the accordance of this proposal with the cultural considerations identified in Chapter Five will be discussed in more detail when exploring the regulative conception of folk psychology and when advancing the adaptive utility of these views below.

this view what constitutes an explanation will be relative to the explanation seeker helps to account for the cultural relativity that is apparent with respect to folk psychological explanation. This relativity allows members of different cultural groups to be correct in their consumption of different explanations of the same behaviour; because the success of an explanation is anchored to a particular individual, this view seamlessly accounts for the plurality of explanations offered by members of different cultures.

As a last point, the role and influence of moral considerations in our folk psychological practice is also consistent with the construal of folk psychological explanation being offered here. Again, because there is no clear restriction on what will resolve an explanation seeker's affective state, there is no *prima facie* inconsistency between the influence of moral considerations and the construction of successful folk explanations. In essence, the asymmetry observed on the basis of the moral valence of the behaviour being explained can play a meaningful role in the construction of explanations so long as this is what is required in order for the explanation seeker to cease seeking an explanation. So long as the explanation offered solves the inquisitor's affective state, then the explanation can be considered an explanation proper and this can include developing explanations that are sensitive to a variety of social, cultural, and moral considerations or more generally the normative and non-causal goals that emerge from the empirical evidence.[279]

This alternative account of folk psychological explanation was already developed in response to an observation regarding how the folk seek and provide explanatory information

[279] There is another important way in which Andrews' account accords with the influence of moral considerations on how we produce and consume folk psychological explanations. As noted, Andrews' view emphasizes the role folk psychology has in defining and maintaining group inclusion criteria. The influence of moral considerations on our folk psychological practice relates to this discussion, but will be discussed below in more detail as it has bearing on the broader discussion of the "alternative" views, and not just Andrews' account.

regarding behaviour, but it is also responsive to the empirical evidence collected and reviewed in this project. In this way, this alternative construal is developed in a way that is aligned with the approach endorsed in Chapter Three; namely, allowing the data to drive the resulting analysis we provide. As such, the proposal here is as plausible as the "multi-purpose" view. However, one advantage of this analysis over the "multi-purpose" view is that this alternative provides a higher level and simpler analysis of the foundation of our folk psychology; that is by supplanting the purpose posited by the traditional construal, this alternative results in a cleaner theoretical analysis than is offered by the "multi-purpose" view.

The extent to which this "alternative" is distinguishable from the "multi-purpose" view is dependent, in part, on whether we take a strong or weak version of this view. I've already noted that this view construes the paradigmatic purpose of folk psychological explanation as the satisfaction of explanation seeking behaviour, but this is not to say that there are not other purposes that shape our explanation seeking, generating, and consuming behaviours. Moreover, this view does not prevent us from building back in a traditional conception of explanation and prediction into the analysis (more on this below). How much emphasis we place on the satisfaction of explanation seeking behaviour as being foundational or primary and whether we can or should situate any other purposes as being secondary to or parasitic on this primary purpose will impact how this "alternative" and the "multi-purpose" view are distinguished. But there does seem to be something unique and fundamentally different about how this alternative treats the *modus operandi* of folk psychology when compared to the "multi-purpose" view. While both views may permit a plurality of interests, this "alternative" view does alter our fundamental understanding of folk psychology in a more dramatic fashion than the "multi-purpose" view.

On this account of folk psychological explanation, we have effectively redefined explanation. So while the construal of folk psychology as an explanatory framework has been retained, what this actually means has changed significantly from the traditional construal. The analysis of explanation offered is broader and more all-encompassing when compared to the narrow view offered by the tradition. In Subsection 6.3.2 above, I noted that this maneuver was available to those trying to retain an explanatory analysis of folk psychology. However, I also noted that reconciling explanation with the social, cultural, and moral analysis would be to the detriment of the traditional construal as we will have had to redefine explanation in such a way that we've retained the term, but lost the core of the traditionalist's view. This is precisely the move that has been made by this alternative. There are two additional points in favour of this view, but as they apply equally well to the second view that is of interest for my purposes here, I'll first turn my attention to that view.

The second view I have identified as being a plausible "alternative" candidate flatly rejects the suggestion that folk psychology is always about explaining and predicting behaviour. Instead, this notion is supplanted with an alternative purpose, namely, the regulation of behaviour. We don't explain and predict behaviour, we shape it. This approach to understanding folk psychology has been taken up by, for example, Mameli (2001), McGeer (2007), and Zawidzki (2008), sometimes under the banner of "mind shaping". On this view, folk psychology permits us to shape the minds of others and in particular, helps to shape our behaviours by articulating and reinforcing the norms of behaviour we target when we act.[280] Again, there is a

[280] There are two suggestions here, one concerning the structure or form of folk psychology and one concerning the function of folk psychology. In terms of the form of folk psychology, the mind shaping account can be committed to the claim that mental states are the paradigmatic cause of behaviour and that folk psychology is concerned primarily with these agent attributes. This claim is consistent with the traditional construal of folk psychology as reviewed in Chapter Two. This is true of, for example, Zawidzki's (2008) contribution to the view. But what is of most interest here is the claim that the mind shaping account makes concerning the function of folk psychology. Specifically, that folk psychology functions to regulate behaviour.

significant departure here from the traditional account insofar as the traditionally construed purpose of folk psychology has been supplanted by a new one. Folk psychology is not just a passive endeavour where it is used to explain or predict behaviour. Instead, we use folk psychology in such a way that it actively shapes behaviour.

On this account, folk psychology is an external framework that regulates behaviour by creating a set of expectations that actors aim to conform to. Mameli (2001) explains this with reference to expectancy effects in psychology. That is, by attributing mental states to ourselves and others we produce an expectancy that the agent act in accordance with this attribution. Generally speaking, "people tend to behave in ways that confirm the beliefs about them held by the persons they interact with" (Mameli, 2001, p.610). Here we see the direct influence of folk psychology on structuring behaviour in regular ways. Supplementing this suggestion, Tadeusz Zawidzki (2008) discusses the role of caretaker expectancies in child development as explored by Dennett (2003, pp.251, 273, 277). In particular, Zawidzki is interested in the observation that "throughout childhood, children are queried for the reasons behind their actions, called on inconsistencies between behavior and utterances, and otherwise prodded to conform to norms of folk psychology, long before they display any competence at such conformity." (Zawidkzi, 2008, p.200) In other words, we find that children are constantly encouraged to think about their behaviours in relationship to the folk psychological norms the parents hold and this "constant prodding" (Zawidzki, 2008, p.200) is likely to be an important factor in child development with respect to seeing behaviour through a folk psychological lens.[281] But importantly, it demonstrates that adult caretakers are (likely unconsciously) directly shaping the minds and behaviour of the

[281] And in particular, Zawidzki notes, through the lens of reasons. I focus, however, more broadly on the folk psychological framework as a whole, as this is just as true as Zawidkzi's focus on reasons but better reflects the commitment throughout this project that folk psychology is more than just beliefs and desires.

children they care for by setting out expectations and norms for behaviour through their folk psychological discourse. This is not just idle speculation as there is evidence to suggest that setting out these norms for children to reflect upon and act in accordance with has a direct effect on our children and the way their behaviour is structured in the long run. For example, Mameli (2001) argues that adults may unconsciously treat infants differently depending on whether they are perceived as male or female and as such, create expectancies in behaviour that later manifest in gender differences. Additionally, Zawidzki notes that caretaker interpretations of early infant vocalizations as being communicative acts may help foster the capacity for intentional communication in infants (Zawidzki, 2008, p.200).[282]

There is also evidence that supports the claim that folk psychology can structure and shape adult behaviour. In particular, Zawidzki reflects on empirical evidence demonstrating that "adults automatically conform to stereotypes with which they are verbally primed." (Zawidzki, 2008, p.200)[283] The automaticity of these adaptations suggests that we are primed to respond to the expectations or norms articulated by others and applied to our own situation.[284]

Ultimately then, what we end up with is an understanding of folk psychology as a framework much like the rules of the road that regulate driving behaviour, where solving

[282] Anecdotally, this accords with my own experience with my son. We started to respond to his familiar "mmmm-ba" vocalization and around 9 months he started to use this sound in a purposeful way when interacting with us.
[283] Zawidzki references a series of studies exploring these effects: Bargh and Chartrand, 1999; Chartrand and Bargh, 1999; Dijksterhuis, 2000.
[284] It is worth noting another couple of examples where verbal utterances shape behaviour; one familiar and speculative and one novel and anecdotal. In the Sellars' Myth of Jones, our Rylean ancestors' behaviour is accompanied by verbal utterances that are articulating the causes or reasons or purpose for behaviour as it is unfolding. This is just another example of our folk psychology actively structuring our behaviour. We also see children and adults employing similar types of techniques when presented with a particularly difficult challenge. For example, a Fifth Grader talking themselves through a long-division problem (e.g. "first I have to divide the number in the hundreds column by the divisor, and then subtract…", etc.) or an adult verbally uttering the procedure required to hook-up their home theatre system (e.g. "this cable has to go to the TV, but first by-pass through the Xbox One before going…", etc.). It is not much of a stretch to see folk psychological utterances as having the same structuring affect, especially as we learn those norms for the first few times. Eventually, of course, these norms become embedded in our practice and effectively disappear.

coordination problems simply depends "on figuring out what the *normatively sanctioned* response to some problem is, and assuming that others will do the same" (Zawidzki, 2008, p.199, original emphasis). Molding behaviour through this "socially instituted normative practice" significantly improves our ability to predict what others are going to do and to coordinate behaviour with others, and through socialization mechanisms we strive to conform to these norms (Zawidzki, 2008, p.199).

Again, while I think this alternative is plausible in its own right, it too can help to make sense of the evidence and conclusions drawn throughout Chapters Four and Five. In fact, this alternative proposal, albeit not intentionally, is actually very responsive to the empirical data collected throughout this project. The social, cultural, and moral considerations that have been identified as driving our folk psychological practice all have something in common: they embody expectations regarding appropriate behaviour. In other words, these considerations are just the embodiment of particular norms regarding behaviour.

While this is true of both social and cultural considerations, it is most obvious when looking at moral considerations identified in Chapter Five. After all, morality is effectively a list of prescriptions that we are all supposed to abide by in order to be a good person and/or to do the right thing. In this way, the considerations that have been identified to shape and drive our folk psychological practice are all regulative in nature; they set expectations regarding how behaviour should unfold. Therefore, insofar as folk psychology is intimately tied up with these social, cultural, and moral considerations it too is regulative and sets expectations for how behaviour should unfold. When we use our folk psychology, we are effectively engaged in an articulation of particular codes of conduct that spell out how we *ought* to act or how behaviour *should* unfold. As such, the regulative account of folk psychology anticipates and expects that folk

psychology has the normative character that has been identified throughout Chapters Four and Five. If folk psychology is a regulative enterprise it is not surprising that we find social, cultural and moral considerations shaping and influencing how we use folk psychological concepts or construct folk psychological explanations and predictions. These considerations are just particular instantiations of the normative structure that the regulative view of folk psychology endorses. While this articulation of the norms that govern behaviour helps to explain why we are often successful in providing accurate explanations and predictions of behaviour (more on this in a moment), it also helps us to make sense of why our folk psychology sometimes fails. It is not as though with every application of folk psychology we are engaging in a deliberative and purposeful investigation of the mental states held by the person we are explaining or the multitude of potential non-propositional or non-mental causes of behaviour. Instead, we are just articulating sets of norms that we hold to generally govern behaviour. This short hand may often be mistaken when the norm misses crucial causally relevant factors or when an individual (intentionally or unintentionally) acts contrary to the norm. But of course our failures are less problematic on this view since the goal is not primarily to explain and predict behaviour. That is, unlike the tradition which claims that our primary goal is to be successful and then needs to address our failures, on this view the primary goal is not the successful explanation and prediction of behaviour, but the regulation and shaping of behaviour.

Whether the regulation of behaviour is the function of folk psychology in all instances is an open question. In fact, Zawidzki notes that there are both a strong and weak version of the regulative account of folk psychology. The strong version maintains that "all of our interpretive practices involve mind shaping" whereas the weaker claim simply states that "s*ome* important aspects of our interpretive practices involve mind shaping." (Zawidzki, 2008, p.194, original

emphasis)[285] To the extent that the weak version is embraced, this particular "alternative" view may actually reduce to the more conservative "multi-purpose" view; regulation just becomes one among the many distinct purposes that folk psychology has.

That said, even the weak version can be distinguished as a significant departure from the traditional construal. In particular, how this view addresses our apparent explanatory and predictive prowess helps to clearly distinguish it from the traditional construal and the other options that have been considered in the preceding subsections. More specifically, how we articulate the relationship between the regulation of behaviour (the alternative purpose) and the explanation and prediction of behaviour (the traditional purpose) may permit us to position the weak version of this alternative as something distinct from the traditional construal and, notably, even the "multi-purpose" view. Recall that the "multi-purpose" view treats these purposes as distinct and co-existent purposes that may often compete. However, on this alternative construal of folk psychology, explanation and prediction can actually be viewed as being parasitic on the regulative nature of folk psychology and as such, significantly different than how the "multi-purpose" view as articulated above would construe the nature and relationship of these purposes.

In brief, if folk psychology is a framework for structuring our behaviour in predictable ways and in ways that conform with social, cultural, and moral expectations regarding behaviour, then our explanatory and predictive success may simply be a result of this fact. For example, if folk psychology is designed in such a way as to embed normatively sanctioned actions and shape behaviour to conform to these norms, then we can predict that people will act in accordance with these norms. Our predictive success is parasitic on the expectations embedded in the framework and the way these expectations structure and shape behaviour. To borrow a phrase from Driver

[285] It is worth noting that Zawidzki adopts the weaker claim.

(2008a), people typically do the morally good thing and so our ability to predict their behaviour relies on this expectation and the norms embedded in our folk psychology.[286] A similar, but somewhat more complicated story, can be offered for explanation. More specifically, our explanatory success stems from an articulation of the same norms that are structuring the observed behaviour. That is, if behaviour tends to conform to a set of norms embedded in our folk psychological discourse, then simply using our discourse will articulate the same norms that have structured the behaviour. In this way, it *looks* like we're producing explanations of behaviour that conform with reality since there is confirmation between the explanation and the behaviour due to their common normative source. As such, whether we adopt a strong or weak version of this alternative account we are speaking about a genuine alternative to the traditional construal of folk psychology, as the traditional purpose is now seen as being derivative of the regulative purpose proposed.

Importantly, a similar maneuver may also be available for Andrews' account of folk psychological explanation, thereby minimizing the likelihood of a potential reduction to the "multi-purpose" view. For example, Andrews notes that anomalous behaviour, the spark that drives explanation seeking behaviour, is only anomalous against the backdrop of normatively appropriate behaviour; we can only make sense of the anomalous with reference to a norm (Andrews, 2012, p.222). In this way, she is similarly open to the notion that our folk psychological discourse is intimately related to the norms of appropriate behaviour. As such, we may be able to adopt a similarly parasitic position with respect to traditional explanation and its relationship to the primary purpose of satisfying explanation seeking behaviour if so desired.

[286] I think this is a fairly robust heuristic. It's not simply a heuristic like "people do what they say they are going to do", but instead, we say that people are going to do what is normatively sanctioned and this is ensured by the fact that the folk psychological framework will be structuring their behaviour.

It is worth emphasizing here that even though both alternatives aim to significantly change our understanding of folk psychology, both are able to help make sense of the intuitive claim that we do go about explaining and predicting behaviour on a regular basis. It's just that on these views, these functions can be seen as being derived from the primary function (i.e. explanation satisfaction or regulation) and in this way the alternatives can be viewed as significantly different from the "multi-purpose" view, but also as the opposite of the "primary purpose view". But this approach does not fall victim to one of the arguments against the "dig in your heels" and "primary purpose" views.

I argued against these views noting that contrary to their claim, the empirical evidence demonstrated that there were some contexts where social, cultural, and moral considerations appeared to adopt a primary role in shaping our folk psychological practices. Moreover, I argued that there were challenges associated with characterizing these normative considerations as being parasitic on the quasi-scientific purposes of explanation and prediction. One might worry that the apparent primacy of the traditional purposes in particular contexts might similarly be a mark against the "alternative" views being explored here. After all, these accounts of folk psychology can be construed as effectively making a claim about the primary purpose of folk psychology that is challenged by the observation that we sometimes have a different primary purpose shaping our practice. However, as I've described above, there is a clear sense in which explanation and prediction, traditionally construed, *can be* parasitic on or derived from the normative framework posited by the "alternative" views. In this way, the "alternative" view can help us make sense of both our successes and failures in a way that is distinguishable from the options explored above.

Finally, we need to make sense of these views from the perspective of adaptive utility. In particular, we must explore how social, cultural, and moral considerations can have adaptive utility that would drive our folk psychological practices to be formed as described by these "alternative" views. While in the subsections above I've hinted at the adaptive utility of these considerations, positioned within these alternative accounts of folk psychology we can begin to provide a robust alternative to the Myth of Jones offered in analogical support of the traditional construal in Chapter Two. A revised and updated Myth of *the Joneses* recognizes that cooperation can serve as the adaptive catalyst to spark and propagate our folk psychological competence. While what follows is true for either alternative being explored in this section, the bulk of the substantive discussion will be anchored to the regulative view for simplicity and ease of developing the argument.

In the *Myth of the Joneses*, we're not interested in just *Jones* as an individual. Rather, we're interested in the *Joneses* as a collective of individuals who comprise a group of our Rylean ancestors. By looking not just at Jones, but at the Joneses we can learn something about folk psychology and how the alternative purposes being discussed in this subsection have significant adaptive utility.

Before looking at the Joneses and the demands of living within a group a bit more closely, let's return to Jones, as described by Sellars, for a moment. There are two things to notice about this myth that help us to make sense of the regulative account of folk psychology and its adaptive utility. First, Sellars may have inadvertently demonstrated the regulative nature of folk psychology in his description of our Rylean ancestors. Recall Sellars' observation that our Rylean ancestors' behaviour is "threaded on overt verbal expressions" (Sellars, 1956/1963, p.186). Presumably, without intention, in this one phrase Sellars actually observes that our folk

psychological discourse is regulative and helps to structure our behaviour. Effectively, Sellars gives an example of how the articulation of folk psychological norms can help to structure one's behaviour in conformity with those norms. As alluded to above (see footnote 284), we sometimes overtly articulate the steps involved in some action in order to structure micro-behaviours in an ordered and appropriate fashion. But Sellars moves from this overt verbal expression to Jones' recognition that this overt behaviour is sometimes not present, but merely occurring within his counterpart's mind. As such, there is a sense in which even within Sellars' original myth, the regulative nature of our folk psychology is primary to the explanatory gap that Jones notices and bridges.

But the primacy of this purpose makes sense when we think a little more about Jones, not as an individual but as a member of a group, call them the Joneses. Even before Jones, the individual, identifies an explanatory gap and then bridges it through the introduction of mentalistic discourse, it's important to notice that the Joneses would have needed to coordinate their behaviour. Put another way, even before Jones developed the ability to "mindread" in order to explain and predict behaviour, there would have been the need for the Joneses to work together and cooperate through coordinated behaviour in order to achieve mutual goals. It would be advantageous for the Joneses to be able to engage in "mindshaping" (to borrow Zawidzki's term) or, more generally, "behaviour-shaping" in order to coordinate behaviour, even before they have the ability to "mindread".

As such, there is a sense in which we could construe the cooperation and coordination of behaviour as being primary to the explanatory and predictive need that Jones, the individual, identified and satisfied. Of course, explanation and prediction might facilitate our ability to cooperate or coordinate behaviour, but the observation here is that the Joneses would have been

coordinating behaviour and cooperating long before Jones' moment of genius and possibly before our need to explain and predict behaviour. But of course, the conception of folk psychology being explored in this subsection is one that facilitates the coordination of behaviour. By prescribing the principles for action, folk psychology structures behaviour in a way that facilitates coordination and cooperation. In this way, the regulative account of folk psychology has adaptive utility. It is of significant utility to be able to achieve ends that would be difficult or impossible to achieve by oneself, and folk psychology as regulation facilitates this achievement and therefore has adaptive utility. Viewing Jones not as an individual, but as a member of the Joneses helps us to focus not on the competitive advantage that folk psychology imparts on the individual, but on the cooperative advantage that folk psychology imparts on the Joneses as a group. Humans are inherently social creatures and it is possible, if not likely, that in an environment where we live and work in close proximity to one another with a need to share a communal space and communal objects that a cooperative approach, not a competitive approach, would be advantageous. We can think of folk psychology as having adaptive utility not because of how it increases just an individual's fitness, but because of how it increases group fitness and thereby the fitness of each individual within the group.

Switching gears, there is another sense in which the regulative nature of folk psychology makes sense from an adaptive perspective. Especially with respect to the claim that successful explanation and prediction are achieved through the attribution of mental states to ourselves and others. This latter approach is actually quite costly. As Zawidzki notes, the "accurate tracking of the incredibly complex neurocognitive properties responsible for hominid behaviour" would require an "improbable accumulation of mutations" (Zawidzki, 2008, p.199) that would take a very long time to evolve randomly. Instead, the regulative account of folk psychology proposes a

much simpler mechanism for shaping our behaviour to make it easier to coordinate behaviour and to predict what others are going to do. Instead of accurately identifying the relevant causal factors and deducing the resulting behaviour, the regulative account of folk psychology simply articulates the normatively sanctioned action in response to one's environment and assumes that others will do the same given that humans have evolved to conform to expectations (Zawidkzi, 2008, p.199). The regulative account of folk psychology offers an approach that makes sense of how we can quickly, easily and accurately make sense of behaviour (Zawidzki, 2008, p.196) that is far less costly, both evolutionarily and cognitively, than devising unique explanations and predictions of observed behaviour in a quasi-scientific way. As such, not only does this understanding of folk psychology make sense from an adaptive utility perspective, but it appears that it is less costly than the story offered in the original Myth of Jones.

But folk psychology as regulation also facilitates the creation and maintenance of groups; something that has significant utility from an evolutionary perspective. Briefly, folk psychology is not something only one person has; it exists within the public space. As a regulative enterprise, folk psychology helps to structure behaviour in a way that conforms to the expectations of an individual's group whether these expectations are social, cultural or moral. As such, folk psychology helps to facilitate group inclusion by structuring behaviour in ways that are sanctioned by the group we belong to. Someone who, for example, consistently acts contrary to our social, cultural or moral principles will be excluded from our group. Moreover, someone who consistently provides explanations and predictions that do not conform to their social, cultural, or moral group's folk psychology would likely not be invited to continue to provide explanations and predictions.

Of course, groups can be defined in a number of different ways or by a number of different considerations. In the language of this project, all of the considerations being explored help to define particular groups. Perhaps most obviously, subscription to a particular set of moral beliefs clearly defines different groups. There may, of course, be overlap between moral groups and some beliefs may cut across a variety of groups, but our morality can help to unite or define a group. Consider for example the group inclusion criteria that are established regarding the ethical treatment of animals and our answers to the question of whether it is ethical to eat meat. Our answers to these questions help to create at least three distinct groups (e.g. meat eaters, vegetarians and vegans). Some groups share ethical beliefs (e.g. that it is ethical to eat the product of an animal), while there are clear divisions between the groups as well (e.g. it is unethical to kill an animal for food). Cultural considerations are also very clearly the basis for group inclusion criteria. Observing the practices, customs, or norms of a particular culture immediately signals to those around you that you are a member of that culture and are committed to that group; a commitment that may encourage others to permit you to share in the success of the group or share in the achievement of a mutually desired goal that would otherwise be difficult to achieve individually. In contrast, not observing those customs or norms immediately signals that you are not a member of that culture and that you are committed to another group; which may be grounds for excluding you from sharing in the goods and efforts of that particular group. Finally, although perhaps less obviously, social considerations can also help to define particular groups by demonstrating a commitment to the norms of a particular social group; for example, political commitments.

The regulative nature of folk psychology then brings about at least three positive consequences. First, it helps to embed our social, cultural, and moral principles in our action

ensuring that those we share a common group with are regularly subject to the same commands. Second, and relatedly, it will also help us identify those with whom we share a common set of norms – both in terms of the explanations and predictions they give and whether they consume our explanations and predictions. Third, it also helps us to structure our behaviour in ways that others will understand because of its conformity to social, cultural, and moral expectations. As there are obviously significant advantages to belonging to a group, by facilitating this process the regulative nature of folk psychology has, again, adaptive utility.

None of this is meant to be a conclusive demonstration of the adaptive utility of the regulative nature of folk psychology. Rather, the goal has been to begin to demonstrate that there is a plausible account of the adaptive utility of this account of folk psychology, even if this story is just in a very preliminary form at this point. Through this investigation we have identified cooperation as the starting point and as being of a potentially great benefit to our survival, the survival of our offspring, and ultimately the species. Viewing folk psychology as a regulative enterprise helps us to make sense of the need we have to cooperate as inherently social creatures and the success cooperation bestows upon us. This success could very likely be the spark that made the creation and continuation of a folk psychological framework that primarily regulates and shapes behaviour possess significant adaptive utility.

Whether we construe the primary purpose of folk psychological explanation as putting an end to the explanation seeking behavior of an explanation seeker or of folk psychology as a whole as predominantly a regulative enterprise, we have an alternative conception of folk psychology at our disposal that is consistent with and makes sense of the empirical evidence explored. While these alternatives construe folk psychology in a way that is consistent with the empirical evidence, they do extend the analysis a little further than the conservative "multi-

purpose" view by proposing to supplant the traditionally construed purpose with a significantly different alternative. While merely offering an adequate analysis of the empirical evidence may not require this extension, the alternatives explored are both plausible views of folk psychology in their own right; or at least they are both deserving of a more comprehensive investigation into their plausibility in addition to their ability to make sense of the empirical evidence. While additional work is required to determine whether these or another "alternative" view is the best analysis of the empirical evidence, this work is again virtuous and productive. The point of this subsection has simply been to identify that there are "alternative" accounts of folk psychology that are significant departures from the traditional view in a way that the "multi-purpose" view is not, but that are also consistent with the empirical evidence explored. In this way, these views are responsive to the empirical evidence and can be built on the back of this work. We can let go of the pause button with a new view of folk psychology in our minds and begin some productive work that is supported by this empirical evidence.

6.4 – Press Play

I began this chapter by arguing that we should press pause and stop to assess our understanding of folk psychology in light of all the empirical evidence explored in the previous chapters. I then explored four options for interpreting the empirical evidence and either preserving the traditional construal of folk psychology or rejecting it as too narrow in scope to properly characterize our folk psychological practices. I advanced a variety of arguments against those options that attempt to preserve the traditional construal and similarly advanced a variety of arguments designed to demonstrate the veracity of the options that reject the tradition and move our understanding of folk psychology in a different direction. While the arguments against and in favour of the options explored were multifarious in nature, the arguments were focused around four key

considerations based on the commitments expressed in Chapters Two and Three and the main findings that emerged from the empirical investigations in Chapters Four and Five.

First, whatever story we present for how the folk engage in folk psychological practices, we must respect the goal of this endeavor, which is to describe the practice as engaged in by the *folk*. We need to recognize that our project is a descriptive one and that our understanding of folk psychology should be based on how the folk engage in these practices, not based on some theoretical commitments that we impose on the folk or on some conception of the practice that accords with philosophical thinking routed in the history of philosophy of action. This commitment was clearly expressed in Chapter Three where I argued that philosophers interested in folk psychology are engaged in a descriptive endeavor and are making empirically testable claims about a real world practice.

The "dig in your heels" and "primary purpose" view do not uphold this commitment. Proponents of these options are effectively preserving their theoretical commitments in light of empirical data that questions the validity of these commitments. What we end up with when exploring these options is a situation much like the resistance that is often expressed before and during scientific revolutions. Consider, for example, Ptolemaic astronomy and the continuous positing of additional orbits to explain the observed planetary movements just to remain committed to the claim that the Earth was the center of the universe. A great deal of extra theorizing was done simply to uphold an assumption that was not being confirmed by the empirical evidence and it took Copernicus to finally toss out this assumption and present a cleaner account of the phenomenon based on the claim that the sun is the center of the universe. Sometimes choosing to save ones' theoretical commitments when there is mounting empirical evidence to the contrary, can be detrimental to the investigative method and may unnecessarily

complicate matters. This is precisely the error that the "dig in your heels" and "primary purpose" views commit. It's time that folk psychology undergoes a Copernican revolution and the "multi-purpose" or "alternative" views offer us two options of how we can go down this path. These two options, I argued, are very much data driven and reflect our observations of how the folk engage in folk psychological practices. Both options express a commitment to the idea that whatever we say about folk psychology, it has to be an accurate description of the practice and both options endeavor to do precisely this.

Once we're committed to this descriptive project, our understanding of folk psychology needs to make sense of the empirical evidence that has been collected and recounted in this project regarding how folk psychology is used in practice by the folk. First, our understanding of folk psychology needs to help us make sense of both our explanatory and predictive successes *and* failures. I've taken seriously the observation that we appear to explain and predict behaviour all the time, in a way that is plausibly consistent with the traditional construal of folk psychology. However, I've also argued on the basis of empirical evidence that we are also error prone in this regard; we are not nearly as successful as the traditional construal of folk psychology would have us believe and it is not appropriate to merely waive these failures off as minor or insignificant in number. Our best understanding of folk psychology needs to make sense of both these observations and as I argued above, the "dig in your heels" and "primary purpose" views appear poorly suited for this purpose. In particular, while these views are perfectly consistent with our successes, they unnecessarily complicate the analysis of our failures. They necessitate that we view the folks' performance as error prone and they must develop an account of why the folk would be error prone in the precise realm that this specialized tool is designed to be used for. As detailed in the subsections above, this complicates

our understanding of the folks' performance and I've argued it's not simply enough to cast this failure as merely a failure in performance. Moreover, it is unnecessary to complicate our analysis of folk psychology in this way when there are other views that can construe this apparent performance failure, as actually successful and as achieving the goal the behaviour set out to achieve. Specifically, the "multi-purpose" view is particularly well suited for accounting for both our successes and our failures and while the "alternative" views explored require that we re-engineer our understanding of why our folk psychological explanations and predictions are successful, they achieve this end while providing a plausible account of why our explanations and predictions fail to meet the standard of the traditional construal.

Most significantly, the empirical evidence discussed in Chapters Four and Five necessitate that our understanding of folk psychology must account for the influence of social, cultural, and moral considerations in shaping our attributions of folk psychological concepts and folk psychological discourse more generally. That is, as practiced by the folk, folk psychology has a normative component whereby our practices are sensitive to and responsive to normative matters; be it social, cultural or moral. This normative quality of folk psychology was apparent through much of Chapters Four and Five, but is perhaps most obvious when examining the influence of moral considerations on how the folk attribute folk psychological concepts to one another the implications this has for how the folk understand agential causation. Any plausible account of folk psychology must make sense of these observations and must be able to account for the normative character of folk psychology that has been observed through our empirical investigations.

It will not, I argued above, suffice to merely characterize the influence of these considerations as distorting or biasing our practices. The performance error argument is not

tenable and we must recognize that these considerations are not merely distortions or biases. They have widespread effects throughout much of our folk psychological practices and in some cases are the primary driver of how the folk use folk psychological concepts or construct descriptions of behaviour. It is precisely because the "dig in your heels" and "primary purpose" views commit these errors that we must reject them. They fail to capture the widespread and precise nature of influence these considerations exert on our folk psychological practice, due to their attempt to effectively explain this influence away. There are, I argued, conceptually simpler and cleaner options available to the investigator of folk psychology that embrace the influence of these considerations as being a genuine effect and as being indicative of a genuine purpose that is distinct from the traditionalist's goal of explanation and prediction. The arguments made above are well-worn at this point, but the "multi-purpose" view recognizes the plurality of goals that can shape folk psychology and accounts for the shifting primacy of goals. The "alternative" views continue to understand folk psychology as being primarily driven by one new goal, what we might call a macro-goal, but one that helps us make sense of the plurality of micro-goals that shape individual acts of folk psychology. Folk psychology can be understood as both descriptive and normative; as a tool for generating explanations and predictions of behavior *and*, more importantly, for facilitating a variety of social, cultural, and moral goals that serve real world purposes.

Lastly, in Chapter Two I emphasized that the traditional construal of folk psychology is plausible, in part, because there is an adaptive story we can tell that helps to make sense of this view. Throughout the discussion of the options I've explored above, I've turned to the question of adaptive utility to explore whether similar stories could be told for each option. Importantly, what we found is that the provision of an adaptive story is not unique to just the traditional

construal, and thus the "dig in your heels" and "primary purpose" view. While these two options allow us to seamlessly attach the analysis of folk psychology to Sellars' original Myth of Jones and the resulting adaptive utility argument that emerged from this myth, it is possible for us to provide similar adaptive utility arguments for the "multi-purpose" and "alternative" views. Insofar as these views account for the explanatory and predictive success of folk psychology, they too can adopt a similar story to Sellars' original myth, but we can also supplant this myth with an alternative story that focuses on the evolution of folk psychology within a social and cooperative environment where there is adaptive utility to the articulation of behavioural norms in our folk psychological discourse.

We're now ready to hit play. I began this Chapter highlighting the incongruity between the empirical evidence and the traditionalist. This led me to press pause and evaluate our options to see if the traditionalist conception of folk psychology can be preserved in light of the evidence. It's now clear that this incongruity warrants more than just skepticism, but an outright rejection of the traditional view. The traditional conception of folk psychology is simply too narrow in scope to adequately account for the empirical evidence and the best options we have available to us require that we reject the traditional conception of folk psychology in favour of a new approach. It's time to hit play, and follow the story into the new folk psychological landscape.

6.5 – The New Landscape and a Conclusion

The story has changed. The traditional construal of folk psychology ultimately casts folk psychology as sharing the same goal of science, namely the explanation and prediction of a natural phenomenon, in this case behaviour, and these explanations and predictions of behaviour are seen to be causal in nature. Chapters Four and Five brought empirical evidence to bear on the

question of whether this construal of folk psychology is accurate. The empirical evidence supported two claims. First that the folk are not as good at explaining and predicting behaviour as they are thought to be by those who endorse the traditional view and second, that a number of social, cultural and moral considerations significantly shape the way our folk psychology is deployed in practice. This evidence, I argued, warranted taking a skeptical position with respect to the traditional construal of folk psychology to assess whether it is in fact an accurate portrayal of our folk psychological practices. This, I argued, is already a significant change in the debate. The traditional view, I noted, has for the most part been accepted as the canon regarding folk psychology. This skeptical position I argued for changed the debate by putting the burden of proof back on the traditionalist, forcing investigators of folk psychology to re-examine and justify some very well entrenched and firmly held assumptions regarding folk psychology.

While this shifting of the burden of proof was in itself significant, we have now gone one step further and rejected the view altogether. I've argued that attempts to save this analysis fail in a number of different ways and that there are alternative conceptions of folk psychology that better account for folk psychological practices as practiced *in situ* by the folk, but that leave the tradition behind. Instead of trying to reconcile the empirical evidence with our theoretical assumptions regarding folk psychology, I have argued that the empirical evidence is such that we must develop a new understanding of how and why the folk use folk psychology. The alternative approaches to understanding folk psychology recognize that the traditional construal is too narrow in scope, and that focusing just narrowly on the goal of quasi-scientific explanation and prediction fails to capture the wide range of distinct and important goals that meaningfully shape our folk psychological practices. Our skepticism has now morphed into a full-scale rejection of the traditional conception of folk psychology in favour of a new landscape of options that

significantly change the way we understand folk psychology and that leave our investigation of folk psychology significantly changed as a result.

Now that we've hit play and are following the story into this new landscape, what can we say about this new setting? While there is much work left to do in order to adjudicate between the different conceptions of folk psychology that are available to us within this new landscape and to articulate the specific character of this new landscape, we've already learned a few things that represent important shifts in our understanding.

In this new landscape, folk psychology is not construed as a specialized tool for explaining and predicting behaviour. Our new conception of folk psychology recognizes that we are not always in the business of providing quasi-scientific explanations and predictions of behaviour when we engage in folk psychological discourse, and as a result we don't always satisfy this goal; we are not as successful at explaining and predicting behaviour as has been assumed by many in the past, in part because this just simply is not always our goal. Instead, there are a multitude of social, cultural, and moral goals that drive these practices. This leads to perhaps the most important conclusion about our understanding of folk psychology in this new landscape. By showing sensitivity to social, cultural, and moral considerations and fulfilling a number of goals in these domains, folk psychology is now to be understood as having a normative dimension. This normative character and the social, cultural, and moral goals that drive our folk psychological practices forces us to confront the assumed relationship between folk psychology and science. In particular, while the traditionalist positions folk psychology as having a scientific goal and being somewhere on the same spectrum as science in terms of methodology, rigour, and approach, our new understanding of folk psychology forces us to question and ultimately reject this claim. Put another way, the traditionalist assumed the answer

to the question “do these theories have the same or similar interests, or are they pursuing different interests?” (Khalidi, 1998, p.44) was that science and folk psychology do in fact share the same interests. The empirical evidence collected throughout this project and the viable conceptions of folk psychology that have been examined and defended, challenge this assumption. Moving forward, we must embrace the alternative answer to this question.

While the traditional construal of folk psychology has assumed that our common sense and scientific endeavours are well aligned, even if roughly so, this need not be the case. Instead, folk psychology and science may share the same target of observation (namely, behaviour), but the underlying function, purpose, or interests that shape these practices need not be the same. This approach to understanding the relationship between common sense and scientific projects is not new (see Dupré, 1981, 1991), and yet the traditional construal of folk psychology has persisted while these types of relationships were contemplated in other domains. Throughout this project, I have demonstrated that there are normative or, in Dupré’s terminology, “human” goals or interests that underlie our folk psychology and these interests are, just as they are for other folk theories, distinct from the scientific endeavour.

For example, just as there are culinary interests (a very human interest) that lead the folk biologist to classify onions and garlic differently than biologists (Dupré, 1981, p.74), there are normative interests that drive a wedge between the relationship that the traditionalist assumes to exist between folk psychology and science. And just as it would be a “severe culinary misfortune if no distinction were drawn between garlic and onions” (Dupré, 1981, p.80) even though this does not accord with our biological understanding of onions and garlic, so to it would be a severe folk psychological misfortune if the normative interests were lost even though they are not consistent with a scientific investigation of behaviour. Our understanding of, for example,

intentional action and person causation is significantly different than our understanding of non-agential causation and interaction; we use folk psychology to describe and understand behaviour in a way that is different than other physical phenomena. And yet, to lose this would be to lose something of significant normative or human value in our folk psychological practices.

Earlier I posed a question. The traditionalist assumes that the difference between folk psychology and science is one of *degree,* but I questioned this assumption in light of all the empirical evidence collected and queried whether it was now necessary to accept that the difference between folk psychology and science is really one of *kind*. In the new landscape we have an answer to this question. We can no longer just accept that the difference between folk psychology and science is merely one of degree. In the new landscape, the difference is a difference in kind.

What we've learned is that our folk psychological practices and in particular, their normative quality, cuts nature at different joints than a quasi-scientific conception of folk psychology would. While within this new landscape we might continue to include a role for explanation and prediction, all this concedes is that there may be some overlap in interests between folk psychology and science. This does not, however, take away from the claim that these approaches are significantly misaligned. At best, it simply concedes that while folk psychology and science may cut nature at different joints, they may also have important intersections or may crosscut one another (see Khalidi (1998) for a discussion of intersecting or crosscutting schemas).

Our understanding of folk psychology has changed. The traditional conception of folk psychology is simply too narrow in scope to adequately account for how the folk deploy folk psychology in practice and the best options we have available to us necessitate a rejection of the

traditional construal in favour of a new landscape. This shift in understanding, however, is not benign. Rejecting the traditionalist conception of folk psychology has potentially broader philosophical implications beyond the domain of just the debates regarding folk psychology. As I noted at the start of this project, folk psychology has played a role in how we understand a number of areas of philosophy. Most notably, it is invoked in metaphysics of mind, the philosophy of artificial intelligence and cognitive science, and even in developmental psychology. By rejecting the traditionalist assumptions about folk psychology it would be a tremendous surprise if this did not have a ripple effect through the other related disciplines. Two potential areas of implication will demonstrate both the significance and diversity of these potential implications.

This shift in understanding regarding folk psychology presents new challenges[287] for those who advance eliminativist or reductionist positions with respect to folk psychology. Briefly, eliminativism (see Churchland, 1981, 1989) and reductionism (see Bickle, 1998, 2003) investigate the relationship between folk psychology and a neuroscientific account of behaviour.[288] The eliminativist argues that folk psychology is so radially false that we ought to reject it outright and replace it with a neuroscientific understanding of behaviour. In contrast, the reductionist argues that folk psychology is not really mistaken, but rather that it should be possible to reduce our folk psychological taxonomy to a neuroscientific one, in effect, translating our common sense conception of behaviour into a neuroscientific one. However, both of these arguments are premised on a claim about the relationship between folk psychology and science.

[287] Or perhaps it is more accurate to say that it "re-presents" challenges for these positions. Dennett (1987) for example, appears to be sensitive to this line of reasoning, but this argument is now bolstered by additional evidence and analysis that was not common in the debate at that time.

[288] I'll refer to the relationship as being between folk psychology and neuroscience, but it need not be just this scientific endeavour. Instead, neuroscience really is just a stand in for any scientific conception of behaviour. The relationship being investigated is most generally one between common sense and scientific theories of behaviour.

In particular, that they have the same object of investigation and that their classifications schemes are appropriately aligned.

The new landscape challenges this premise as in this new setting, at best folk psychological and scientific classifications intersect or crosscut one another, but fundamentally, they differ in kind. As such, it is possible that within this new landscape eliminativism and reductionism may not be tenable positions. Worse, even if it could somehow be argued that there is an alignment of the relevant interests that are necessary to allow for the potential elimination or reduction of folk psychology, I would worry that the normative character of folk psychology identified throughout this project would make this move inadvisable. This normative character may not be reducible and eliminating this normative character would be to lose a significant human interest in our folk psychological practices.

There could also be implications for how we understand the relationship between moral psychology and folk psychology and this new landscape may present additional practical challenges for our judicial system and the legal theorizing that justifies it. Consider first, as an example, the Doctrine of Double Effect (see Foot, 1967). This model for moral judgment making is premised on our ability to *first* asses an actor's intentions with respect to a side-effect prior to making a judgment regarding whether their action is worthy of moral condemnation; folk psychological judgments *inform* moral judgments. But of course, the new landscape accepts that moral judgments can and do inform folk psychological judgments. In practice, it appears that moral psychology and folk psychology work in tandem or in a symbiotic relationship rather than in a neat and unidirectional manner and the impact of this multidirectional relationship warrants consideration as we develop theories of moral psychology.

But this new relationship between moral and folk psychological judgments also creates practical difficulties for how we envision ideal decision-making in the legal system.[289] In particular, the judicial system in many countries is built upon the presumption of innocence in order to avoid biasing any decision-making that ensues regarding the guilt or innocence of the charged. As a part of this process, assessments of the defendant's causal responsibility, beliefs, desires, knowledge and intentions will all likely need to be made which then inform assessments about moral or legal culpability. But inherent in our new understanding of folk psychology is the claim that we are not well equipped to separate these questions out. As we have learned the mere fact that an outcome is morally objectionable influences attributions of a wide range of folk psychological concepts including causal responsibility. Again, that the interplay between folk psychological and moral judgments is not unidirectional will further complicate judiciary decision-making processes.

There is much left to do. After over thirty years of operating within the traditional construal of folk psychology, it may seem foreign to depart from this view and may seem daunting to begin to undertake work within this new landscape. But this work is virtuous and productive and is required if we are going to be genuine investigators of folk psychology and remain committed to our goal of describing the practice as practiced by the folk. I've identified two options for how we can understand folk psychology in this new landscape, but much work remains to develop and refine these conceptions of folk psychology and adjudicate between them. Moreover, we must be open to other conceptions of folk psychology not explored here that may also account for the empirical evidence analyzed throughout this project. We must also

[289] Thank you to Paul J. I. Alexander, J.D. for suggesting this to me in a discussion of the empirical results found in Roxborough and Cumby (2009). I have built upon his suggestion taking into consideration the broader folk psychological framework (see also, Nadelhoffer, 2006).

recognize that our foray into empirical investigations of folk psychology are not yet complete. There is likely significant value in continuing to identify and understand all the various interests that work together to shape and drive our folk psychological practices so that we can ensure our understanding of these practices is as empirically informed as possible.

Notwithstanding all the work that remains and that there is still much to learn about this new landscape, we already know a few important things that represent significant changes in our understanding of folk psychology. We know that at a minimum, we must embrace the idea that we are not always as good at explaining and prediction behaviour as has been assumed in the past and that there are a host of normative interests, be they social, cultural, or moral, that shape and drive this practice and we must accept that the relationship between science and folk psychology has been changed in a significant way. This new landscape is a significant departure from the traditionalist conception of folk psychology that we explored at the beginning of this project. Where we go from here is still an open question, but wherever the story takes us one thing is clear: we've put the *folk* back into folk psychology.

Bibliography

Adams, F., & Steadman, A. (2004a). Intentional action in ordinary language: Core concept or pragmatic understanding? *Analysis, 64*, 173-181.

---------. (2004b). Intentional action and moral considerations: Still pragmatic. *Analysis, 64*, 268-276.

Andrews, K. (2003). Knowing mental states: The asymmetry of psychological prediction and explanation. In Q. Smith & A. Jokic (Eds.), *Consciousness: New Philosophical Perspectives* (pp. 201-219), Oxford, UK: Oxford University Press.

---------. (2008). It's in your nature: A pluralistic folk psychology. *Synthese, 165*(1), 13-29.

---------. (2012). *Can apes read minds? Toward a new folk psychology*. Cambridge, MA: MIT Press.

Aristotle. *The Nicomachean Ethics*. (W.D. Ross, Trans.)

---------. *De Anima*. (J.A. Smith, Trans.)

Bargh, J.A., & Chartrand, T.L. (1999). The unbearable automaticity of being. *American Psychologist, 54*, 462-479.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind*. Cambridge, MA: MIT Press.

Bayne, T. (2011). The presence of consciousness in absence seizures. *Behavioural Neurology, 24*, 47-53.

---------. (2008). The unity of consciousness and the split-brain syndrome. *Journal of Philosophy, 105*(6), 277-300.

Bealer, G. (1998). Intuition and the autonomy of philosophy. In M. De Paul & W. Ramsey (Eds.), *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 201-240), Lanham, MD: Rowan & Littlefield.

Bickle, J. (1998). *Psychoneural reduction: The new wave*. Cambridge, MA: MIT Press.

---------. (2003). *Philosophy and neuroscience: A ruthlessly reductive account*. Dordrecht, NL: Kluwer Academic Publishers.

Block, N. (1978). Troubles with functionalism. *Minnesota Studies in the Philosophy of Science, 9*, 261-325.

---------. (2004). Qualia. In R. Gregory (Ed.), *Oxford companion to the mind* (2nd ed., pp. 785-789), New York, NY: Oxford University Press.

Borgida. E., & Nisbett, R.E. (1977). The differential impact of abstract vs. concrete information on decisions. *Journal of Applied Social Psychology, 7*(3), 258-271.

Brassolotto, J., Raphael, D., Baldeo, N. (2014). Epistemological barriers to addressing the social determinants of health among public health professionals in Ontario, Canada: a qualitative inquiry. *Critical Public Health*, *24*(3), 321-336.

Buckwalter, W. & Stich, S. (2010). Gender and philosophical intuition. Available at SSRN: http://ssrn.com/abstract=1683066

Cappelan, H. (2012). *Philosophy without intuitions*. Oxford, UK: Oxford University Press.

Carruthers, P. (2011). *The opacity of mind: An integrative theory of self-knowledge*. Oxford, UK: Oxford University Press,

Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200-219.

Chartrand, T.L., & Bargh, J.A. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893-910.

Christensen, S.M., & Turner, D.R. (1993). *Folk psychology and the philosophy of mind*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Churchland, P.M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy, 78*, 67-90.

---------. (1989). Folk psychology and the explanation of human behavior. *Philosophical Perspectives, 3*, 225-241.

Churchland, P.S. (1988). Reduction and the neurobiological basis of consciousness. In A.J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 273-304). Oxford, UK: Oxford University Press

Cushman, F.A. (2010). The effect of moral judgment on causal and intentional attribution: What we say, or how we think? Unpublished manuscript.

Cushman, F.A., & Mele, A. (2008). Intentional action: Two-and-a-half folk concepts? In J. Knobe and S. Nichols (Eds.), *Experimental philosophy* (pp. 171-188). New York, NY: Oxford University Press.

Cushman, F.A., Young, L., & Hauser, M.D. (2006). The role of conscious reasoning and intuitions in moral judgment: Testing three principles of harm. *Psychological Science, 17*(12), 1082-1089.

Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy, 60*(23), 685-700.

Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press.

---------. (2005). *Sweet dreams: Philosophical obstacles to a science of consciousness*. New York, NY: Basic Books.

Dijksterhuis, A. (2000). Behavioral indecision: Effects of self focus on automatic behavior. *Social Cognition, 18*, 55–74.

Dretske, F. (1970). Epistemic operators. *The Journal of Philosophy, 67*(24), 1007-1023.

Driver, J. (2008a). Attributions of causation and moral responsibility. In W. Sinnott-Armstrong (Ed.), *Moral psychology (Vol: 2): The cognitive science of morality: Intuition and diversity* (pp. 423-439), Cambridge, MA: MIT Press.

---------. (2008b). Kinds of norms and legal causation: Reply to Knobe and Fraser and Deigh. In W. Sinnott-Armstrong (Ed.), *Moral psychology (Vol: 2): The cognitive science of morality: Intuition and diversity* (pp. 459-461). Cambridge, MA: MIT Press.

Dupré, J. (1981). Natural kinds and biological taxa. *Philosophical Review, 90*, 66-90.

---------. (1999). Are whales fish? In D.L. Medin & S. Atran (Eds.), *Folkbiology* (pp. 461-476), Cambridge, MA: MIT Press.

Eliasmith, C. (2003). Moving beyond metaphors: Understanding the mind for what it is. *The Journal of Philosophy, C*(10), 493-520.

Fodor, J. (1987). *Psychosemantics: The problem of meaning in the philosophy of mind*. Cambridge, MA: MIT Press.

---------. (1991). You can fool some of the people all of the time, everything else being equal; Hedged laws and psychological explanation. *Mind, 100*, 19-34.

Foley, R. (1998). Rationality and intellectual self-trust. In M. De Paul & W. Ramsey (Eds.) *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 241-256), Lanham, MD: Rowan & Littlefield.

Foot, P. (1967). The problem of abortion and the doctrine of the double effect in virtues and vices, *Oxford Review, 5*, 1-7.

Fox, K.C.R., Zakarauskas, P., Dixon, M.L., Ellamil, M., Thompson, E. & Christoff, K. (2012). Meditation experience predicts introspective accuracy. *PLoS ONE, 7*(9), e45370.

Gallagher, S. (2001). The practice of mind: Theory, simulation, or interaction? *Journal of Consciousness Studies, 8*(5–7), 83-107.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mindreading. *Trends in Cognitive Sciences, 2*, 493-501.

Gettier, E. (1963). Is justified true belief knowledge? *Analysis, 23*, 121-123.

Goldman, A. (1986). *Epistemology and cognition*. Cambridge, MA: MIT Press.

---------. (1993). Consciousness, folk psychology, and cognitive science. *Consciousness and Cognition, 2*, 364-382.

---------. (1995). Interpretation psychologized. In M. Davies & T. Stone (Eds.) *Folk psychology* (pp. 74-99), Oxford, UK: Blackwell.

---------. (2006). *Simulating minds: The philosophy, psychology and neuroscience of mindreading*. New York, NY: Oxford University Press.

Gopnik, A., & Meltzoff, A. (1997). *Words, thoughts and theories*. Cambridge, MA: MIT Press.

Gopnik, A., & Wellman, H. (1992). Why the child's theory of mind really is a theory. *Mind and Language, 7*, 145-171.

Gordon, R. (1986). Folk psychology as simulation. *Mind and Language, 1*, 158-171.

---------. (1995). The simulation theory: Objections and misconceptions. In M. Davies & T. Stone (Eds.), *Folk psychology* (pp. 100-122), Oxford, UK: Blackwell.

Graham, G. (1987). The origins of folk psychology. *Inquiry, 30*(4), 357-379.

Graham, G., & Horgan, T. (1988). How to be realistic about folk psychology. *Philosophical Psychology, 1*(1), 69-81.

Guglielmo, S., & Malle, B.F. (2008). Can unintended side-effects be intentional? Solving a puzzle in people's judgments of intentionality and morality. *Personality and Social Psychology Bulletin, 36*(12), 1635-1647.

Haidt, J., Koller, S., & Dias, M. (1993). Affect, culture and morality. *Journal of Personality and Social Psychology, 65*(4), 613-628.

Hales, S. (2003). Can science debunk rational intuition? Presented at 21st World Congress of Philosophy, Istanbul, Turkey.

Hart, H.L.A., & Honoré, A.M., (1959). *Causation in the law*. Oxford, UK: Clarendon Press.

Heil, J. (2004). *Philosophy of mind: A contemporary introduction*, Second Edition. New York, NY: Routledge.

Hitchcock, C., & Knobe, J. (2009). Cause and norm. *Journal of Philosophy, 106*(11), 587–612.

Hume, D. (1739/1896). *A treatise of human nature*. L.A. Selby-Bigge (Ed.), Oxford, UK: Clarendon Press.

Humprey, N. (1986). *The inner eye*. Oxford, UK: Oxford University Press.

Hutto, D.D. (2008). *Folk psychological narratives: The sociocultural basis of understanding reasons*. Cambridge, MA: MIT Press.

Jones, E.E., & Nisbett, R.E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E.E. Jones, D. Kanouse, H.H. Kelley, R.E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79-94), Morristown, NJ: General Learning Press.

Khalidi, M.A. (1998). Natural kinds and crosscutting categories. *Journal of Philosophy, 95*, 33-50.

Kim, J. (2006). *Philosophy of mind*. Second Edition. USA: Westview Press.

Knobe, J. (2003). Intentional action and side effects in ordinary language. *Analysis, 63*(3), 190-194.

---------. (2004). Intention, intentional action and moral considerations. *Analysis, 64*(2), 181-187.

---------. (2006a). *Folk psychology, folk morality*. Unpublished doctoral dissertation. Princeton, University.

---------. (2006b). The concept of intentional action: A case study in the uses of folk psychology. *Philosophical Studies, 130*(2), 203-231.

Knobe, J., & Burra, A. (2006). The folk concepts of intention and intentional action: A crosscultural study. *Journal of Culture and Cognition, 6*, 113-132.

Knobe, J., & Fraser, B. (2008). Causal judgment and moral judgment: Two experiments. In W. Sinnott-Armstrong (Ed.), *Moral psychology (Vol: 2): The cognitive science of morality: Intuition and diversity* (pp. 441-447), Cambridge, MA: MIT Press.

Knobe, J., & Malle, B.F. (2002). Self and other in the explanation of behavior: 30 years later. *Psychologica Belgica, 42*(1-2), 113-130.

Kuhn, T. (1962). *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press.

Ladyman, J., & Ross, D. (2007). *Every thing must go*. Oxford, UK: Oxford University Press.

Lehrer, K. (1990). *Theory of knowledge*. Boulder, CO: Westview Press and Routledge.

Leslie, A. (1987). Pretense and representation: the origins of "theory of mind". *Psychological Review, 94*, 412-426.

Leslie, A., Knobe, J., & Cohen, A. (2006). Acting intentionally and the side-effect effect: 'theory of mind' and moral judgment. *Psychological Science, 17*, 421-427.

Lewis, D. (1972). Psychophysical and theoretical identifications. *Australasian Journal of Philosophy, 50*, 249-258.

Livengood, J., Sytsma, J., & Rose, D. (2011). Following the FAD: Folk attributions and theories of actual causation. Unpublished manuscript.

Lord, C.G., Ross, L., & Lepper, M.R. (1979). Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence. *Journal of Personality and Social Psychology, 37*(11), 2098-2109.

Lowe, E.J. (2000). *An Introduction to the philosophy of mind*. Cambridge, UK: Cambridge University Press.

Ludwig, K. (2007). The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy, 31*(1), 128-159.

Machery, E. (2008). The folk concept of intentional action: philosophical and experimental issues. *Mind & Language, 23*, 165–89.

Machery, E., Mallon, R., Nichols, S., & Stich, S. (2004). Semantics, cross-cultural style. *Cognition, 92*, B1-B12.

Maibom, H. (2003). In defence of (model) theory theory. *Journal of Consciousness Studies, 16*(6-8), 360-378.

---------. (2005). Moral unreason: The case of psychopathy. *Mind & Language, 20*(2), 237-257.

Malle, B.F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: Bradford Book/MIT Press

Mameli, M. (2001) Mindreading, mindshaping and evolution. *Biology and Philosophy, 16*, 597-628.

McCann, H. (2005). Intentional action and intending: recent empirical studies. *Philosophical Psychology, 18*, 737-748.

McGeer, V. (2007). The regulative dimension of folk psychology. In D. Hutto and M. Ratcliffe (Eds.), *Folk psychology re-assessed* (pp. 138-156), Dordrecht, NL: Springer.

Mele, A.R., & Sverdlik, S. (1996). Intention, intentional action, and moral responsibility. *Philosophical Studies, 82*, 265-287.

Miller, J.G. (1984). Culture and the development of everyday social explanation. Journal of *Personality and Social Psychology, 46*, 961-978.

Morris, M.W., & Larrick, R.P. (1995). When one cause casts doubt on another: A normative analysis of discounting in causal attribution. *Psychological Review, 102*(2), 331-335.

Nadelhoffer, T. (2006). Bad acts, blameworthy agents, and intentional actions: Some problems for jury impartiality. *Philosophical Explorations, 9*, 203-220.

Nado, J. (2008). Effects of moral cognition on judgments of intentionality. *British Journal for the Philosophy of Science, 59*, 709-731.

Nichols, S., & Stich, S. (2003). *Mindreading: An integrated account of pretence, self-awareness, and understanding other minds*. Oxford, UK: Oxford University Press.

Nichols, S., & Ulatowski, J. (2007). Intuitions and individual differences: The Knobe effect revisited. *Mind & Language, 22*, 346-365.

Nisbett, R.E. (2003). *The geography of thought: How Asians and Westerners think differently and why*. New York, NY: Free Press.

Nisbett, R.E., Borgida, E., Crandall, R., & Reed, H. (1976). Popular induction: Information is not always informative. In J. Carrol & J. Payne (Eds.), *Cognition and social behavior* (pp. 227–236), Hillsdale, NJ: Lawrence Erlbaum.

Nisbett, R.E., Peng, K., Choi, I., & Norenzayan, A. (2001). Culture and systems of thought: Holistic vs. analytic cognition. *Psychological Review, 108*, 291-310.

Nisbett, R.E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.

Nisbett, R.E., & Wilson, T. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231-259.

Norenzayan, A., Nisbett, R.E., Smith, E.E., & Kim, B.J. (1999). *Rules vs. similarity as a basis for reasoning and judgment in East and West*. Ann Arbor, MI: University of Michigan.

Onishi, K.H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308(5719),* 255-258.

Pettit, D., & Knobe, J. (2009). The pervasive impact of moral judgment. *Mind & Language, 24*, 586-604.

Pollock, J., & Cruz, J. (1999). *Contemporary theories of knowledge*. Lanham, MD: Rowman and Littlefield.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioural and Brain Sciences, 25*, 1-20.

Quartz, S. (1999). The constructivist brain. *Trends in Cognitive Sciences, 3*(2), 48-57.

Ramsey, W. (1992/1998). Prototypes and Conceptual Analysis. In M. De Paul & W. Ramsey (Eds.), *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 161-178), Lanham, MD: Rowan & Littlefield.

Roe, A., Pallas, S., Hahn, J., & Sur, M. (1990). A map of visual space induced in auditory cortex. *Science, 250,* 818-820.

Roskies, A.L. (2002). Neuroethics for the new millennium. *Neuron, 35*, 21-23.

Roxborough, C., & Cumby, J. (2009). Folk psychological concepts: Causation. *Philosophical Psychology, 22*(2), 205-213.

Roxborough, C., Cumby, J., & Fraser, B. (2010). Causation, morality & atypicality. Unpublished manuscript.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition, 49*, 11-36.

Schlagger, B., & O'Leary, D. (1991). Potential of visual cortex to develop an array of functional units unique to somatosensory cortex. *Science, 252*, 1556-1560.

Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York, NY: HarperCollins Publishers

Schwitzgebel, E. (2009). Do ethicists steal more books? *Philosophical Psychology, 22*, 711-725.

Schwitzgebel, E., & Cushman, F.A. (2012). Expertise in moral reasoning? Order effects on moral judgment in professional philosophers and non-philosophers. *Mind & Language, 27*, 135-153.

Searle, J. (1980). Minds, brains, and programs. *Behaviour and Brain Sciences, 1*, 417-424.

---------. (1994). *The rediscovery of mind*. Cambridge, MA: MIT Press.

Sellars, W. (1956/1963). Empiricism and the philosophy of mind. In *Science, perception and reality* (pp. 127-196). London, UK: Routledge & Kegan Paul.

Stich, S. (1988/1998). Reflective equilibrium, analytic epistemology and the problem of cognitive diversity. In M. De Paul & W. Ramsey (Eds.), *Rethinking intuition: The psychology of intuition and its role in philosophical inquiry* (pp. 95-112), Lanham, MD: Rowan & Littlefield.

---------. (1990). *The fragmentation of reason: Preface to a pragmatic theory of cognitive evaluation*. Cambridge, MA: MIT Press.

Stich, S., & Ravenscroft, I. (1994/1996). What is folk psychology? In S. Stich, *Deconstructing the mind* (pp. 115-135), Oxford, UK: Oxford University Press.

Sytsma, J., Livengood, J., & Rose, D. (2012). Two types of typicality: Rethinking the role of statistical typicality in ordinary causal attributions. *Studies in History and Philosophy of Biological and Biomedical Sciences, 43*, 814-820.

Sytsma, J., & Machery, E. (2010). Two conceptions of subjective experience. *Philosophical Studies, 151*(2), 299-327.

Swain, S., Alexander, J., & Weinberg, J.M. (2008). The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research, 76*(1), 138-155.

Tannenbaum, D., Ditto, P. H. & Pizarro, D.A. (2009) "Different moral values produce different judgments of intentional action". Unpublished manuscript.

Tiles, M. (1984). *Bachelard: Science and objectivity*. Cambridge, UK: Cambridge University Press.

Tversky, A., & Kahneman, D. (1973) Availability: A heuristic for judging frequency and probability. *Cognitive Psychology, 5*, 207-232.

von Eckardt, B. (1994). Folk psychology. In S. Guttenplan (Ed.), *A companion to the philosophy of mind* (pp. 300-307), Malden, MA: Blackwell Publishing.

Weinberg, J.M. (2007). How to challenge intuitions empirically without risking skepticism. *Midwest Studies in Philosophy, 31*(1), 318-343.

Weinberg, J.M., Nichols, S., & Stich, S. (2001). Normativity and epistemic intuitions. *Philosophical Topics, 29*(1&2), 429-460.

Williamson, T. (2007). *The philosophy of philosophy*. Oxford, UK: Blackwell.

Wilson, T.D., Lisle, D.J., Schooler, J.S., Hodges, S.D., Klaren, K.J., & LaFleur, S.J. (1993). Introspecting about reasons can reduce post-choice satisfaction. Personality and Social *Psychology Bulletin, 19,* 331-339.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*, 103-128.

Wright, J., & Bengson, J. (2009) Asymmetries in judgments of responsibility and intentional action. *Mind & Language, 24*, 24-50.

Young, L., Cushman, F.A., Adolphs, R., Tranel, D. & Hauser, M. (2006). Does emotion mediate the effect of an action's moral status on its intentional status? Neuropsychological evidence. *Journal of Cognition and Culture, 6*, 265-278.

Zawidzki, T. (2008). The function of folk psychology: Mind reading or mind shaping? *Philosophical Explorations, 11*(3), 193-210.