

# Infallibilism And the State of Nature

An Mphil Thesis by Donald Berry

University College London

## Signed Declaration

I, Don Berry confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

## Abstract

The subject of this thesis is the debate between infallibilism and fallibilism in epistemology. Rather than the usual approach of trying to encapsulate intuitions regarding the meaning of the term ‘knowledge’ with a set of logically necessary and sufficient conditions, I have instead – following Edward Craig’s book ‘Knowledge and State of Nature’ (Craig, 1990) – begun with an hypothesis about the role of the concept in everyday life, and used this to construct a picture of the prototypical case of ascription when this purpose is properly fulfilled. The hypothesis employed is that the concept is used to ‘flag good informants’, and the primary methodological technique employed is consideration of the situation of an individual inquirer in the state of nature. Later on (in Chapter 2), considerations brought in by the fact that life in a primitive community is inherently social are brought in, thus complicating the picture; the effect is to derelativize it from the individual situation of the enquirer and bring it closer to the familiar analyses of the literature (with which I make various comparisons). I use the picture of the concept of knowledge thus constructed as a basis to argue against the infallibilists’ positions in chapter 3, and construct a positive alternative in chapter 4.

## Contents

Chapter 1: ‘The State of Nature’	p3
Chapter 2: ‘The Concept of Knowledge’	p14
Chapter 3: ‘Some Varieties of Epistemic Infallibilism’	p28
Chapter 4: ‘Interest-Relative Contextualism’	p42

# 1. The State of Nature

## 1.1 Introduction: Nature and Purpose of Project

This thesis will aim to contribute to a debate in Epistemology, namely that between fallibilism and infallibilism. In this thesis I will argue that infallibilism about knowledge is false – that is, knowledge ascriptions can be made correctly whilst the infallibilist's conditions are not met.

The structure of the thesis will proceed as follows. In this chapter, I will propose a methodology under which we can approach the problem, and various other problems along the way: this is, with a few modifications, the strategy of 'conceptual synthesis' found in Edward Craig's book 'Knowledge and the state of nature' (Craig, 1990). This approach proceeds by first hypothesizing a role for the *concept* of knowledge, and then investigating the features we should expect a concept suited to this role to have. In the second chapter I will begin carrying out this investigation, arriving at a broad outline of the concept of knowledge, and then categorizing it according to distinctions such as internalist/externalist.

In the third chapter I will use the Craigean framework to argue against infallibilism about knowledge. There are various formulations of infallibilism I shall consider; Unger's non-accidentalism, the exclusion principle, the 'conclusive reasons' formulation found in Dretske. In the fourth and final chapter I will explore how the Craigean position fits in with recent developments in the literature – referring to the work of Jason Stanley – which claim that facts about practical rationality must be incorporated into the truth conditions of knowledge ascriptions.

My thesis may thus be seen as complimentary to Craig's project: I aim to extend his ideas to a particular problem that he discusses in his book only briefly – that of whether or not knowers must be in some sense infallible – and further develop his position by forging a connection between his work and subsequently published literature.

In the opening pages of his book 'Knowledge and the state of nature', Edward Craig describes the most commonly addressed problem in modern epistemology – that of providing an analysis of the meaning of the term 'know' and its cognates. The standard approach in attempting to answer it is to try to characterise the term's extension, as ascertained by exploring our 'intuitions', with an explicit intention couched in terms of necessary and sufficient conditions (p1). However, he claims that this project is deficient in at least two respects.

The first respect in which it is deficient is that we not only have intuitions about the term's extension, but also intuitions that relate more directly to its intension: hence in a given case, we might not only feel that a subject is or is not in possession of knowledge, but also have an idea of *why* it does or does not qualify (because the subject came about their beliefs by accident, for example). Craig goes on; 'The sceptic notoriously tries to show that the two do not mesh: our intuitions about the intension, the conditions of application of the concept, in fact determine a much smaller extension than that which our directly extensional intuitions mark out' (p1). If the two sets of intuitions disagree then it is not clear how to progress, and philosophers will swiftly become divided into

opposing camps. Moreover, this is even supposing that there is such a thing as a stable set of intuitions about the extension (or intension) of the concept that do not vary from person to person, or with the context in which an umpire is asked.

The second respect in which this project is deficit, or at least limited in scope, is that even supposing the former line of enquiry were brought to a conclusion there would still be a puzzle remaining: why is it a term whose meaning is given by precisely *these* conditions that we find worth bothering to demarcate and talk about? Even supposing (I shall have something to say about this in the next section) that knowledge is something of a 'natural kind', so that its demarcation would not be arbitrary, we might still wonder why it is that knowledge thus understood plays such an important role in our lives.

An alternative approach to this aim of analysing the term 'knowledge', one which Craig suggests will be more promising, is to begin not with an intuitive extension but rather an hypothesis about the role of the *concept* of knowledge: how it is used in everyday life, in pursuit of certain very general practical needs. Given this hypothesis, we may then explore the properties a concept suited to this role would have, and proceed to an analysis of the phenomenon of knowledge itself. We will thus have another criterion of success: as well as fitting at least reasonably well with both our extensional and intensional intuitions, the concept we construct should be fit for use in this particular role. Moreover, such a project would clearly be far better suited to tackling our second question about the importance of knowledge: at least part of the reason it finds such an important role in our lives must be its usefulness in answering to these needs. Craig goes on: 'Such an investigation would still have an anchorage point in the everyday concept: should it reach a result quite different from the intuitive extension, then, barring some special and especially plausible explanation of the mismatch, the original hypothesis about the role that the concept plays in our life would of course be the first casualty' (Craig, 1990, p.2), whence we would need to start from scratch with a different hypothesis.

Because of this emphasis on the role of the concept in everyday life, the approach thus described has strong connections to the tradition of naturalism, which is described by Craig as one in which 'thinkers see man, his behaviour and institutions, as natural facts to be understood as the (broadly speaking causal) outcome of other natural facts... What concepts we use, what linguistic practices are common amongst us, these are special cases of input to the more general naturalistic enterprise' (Craig, 1990, p9). The hypothesis Craig suggests, then, is that the purpose of the concept of knowledge is to 'flag good informants' regarding anything we might be (usually practically) interested in finding out about. I shall argue that if this hypothesis is correct, then knowledge is best seen not as something absolute and rarely attained, but rather an important social institution; a commonplace relation to which we all stand in with respect a large number of everyday propositions.

## 1.2 Objections to the Conceptual Strategy

Before elaborating the method in more detail, I will pause here to anticipate two potential objections. The first is: why should we concern ourselves with the concept of knowledge,

rather than investigating the phenomena of knowledge itself? As Craig points out (p3)<sup>1</sup>, there are situations where the conceptual strategy will not yield any interesting results: for instance, the point of the concept of water is to enable us to talk and think about water, and no amount of analysis of the role of this concept will allow us to learn anything new about water itself. In response, Craig suggests that ‘Knowledge is not a given phenomenon, but something that we delineate by operating with a concept which we create in answer to certain needs, or in pursuit of certain ideals’ (p3). Thus the limits of knowledge do not mark the transition from one objectively distinct kind of phenomena to another, but are rather drawn in place in response to certain practical needs: in other words, knowledge is not a ‘natural kind’, and it is the concept and associated conceptual practises that are primary. Hence in order to understand the properties that knowers must possess and why, we must look to the forces that govern our conceptual practices (which we may take to be the actual application of the term in day to day life).

How might we argue for this claim? Rather than arguing for it directly, Craig instead claims that the ‘proof is in the pudding’: he suggests proceeding with his experiment, and resting his case upon the explanatory power of the analysis that he develops, and the light it sheds on debates within the literature that had otherwise seemed stalled.

Although pursuit of this idea would take us two far afield, there is another line of thought that might be worth following. Later I shall argue that belief is necessary for knowledge, and many thinkers have claimed that belief is not a natural kind. For example, in the intentional psychology of Daniel C. Dennett (Dennett, 1989, p15) we find an ‘instrumentalist’ attitude towards belief: although beliefs do corresponding to objectively existing patterns, we should not expect to find individually distinguishable elements in the brain corresponding to distinct beliefs; beliefs and other intentional entities are seen as primarily theoretical constructs that derive from concepts which are projected onto individuals in order to make sense of and predict their behaviour (we adopt the ‘intentional stance’ towards them). If beliefs do not constitute a natural kind, and knowledge is intimately connected to belief, then it becomes plausible to suggest that knowledge is not a natural kind either.

A second objection to Craig’s proposal is that some philosophers may have concerns about the prospective value of this endeavour: why should we be interested in everyday conceptual practise at all? ‘Surely’, a critic might respond, ‘there is a difference between what knowledge is, and what people *say* it is. If our only aim is to describe popular, everyday usage of the term, then we have ceased to even be doing Philosophy.’ The thought is that although this kind of project might be interesting to a linguist or a sociologist, a philosophical analysis of knowledge should go beyond everyday usage and aim at an analysis of knowledge proper. As Philosophers we should provide a prescriptive rather than a merely descriptive account, because our particular set of skills leave us better placed to determine its nature and proper application than members of the population at large.

My response to this concern (which again draws on Craig’s ideas) is twofold. Firstly, we are not merely aiming at a description of certain linguistic facts, but also seeking to clarify certain aspects of conceptual practice: we can analyse the *proper* deployment of a concept suited to the role we have hypothesized, not the *de facto* way in which it is

---

<sup>1</sup> Where not specified, references are to Craig, 1990.

currently deployed. Thus the prescriptive element is introduced at a different place. If we were only interested in linguistic usage, the question would be trivial: it is clear that people in general do not make the demands the infallibilist insists they should when ascribing knowledge. As we shall see, there are three levels of analysis by which we can proceed, and this objection is only really appropriate to one of them (the ‘descriptive framework’ – see section 1.4 below).

Secondly, if, as we have claimed, knowledge is not a natural kind but rather a phenomena delineated by our conceptual practises, then constructing a concept suitable for a particular role may be the best kind of analysis we can hope for. It might be true that the purpose we now need the concept for is different from the one in relation to which it arose, but by grasping the connection between the everyday concept as it stands and its initial purpose we are better placed to make any adjustments that might be necessary to tailor it to be suitable for our current concerns (such as deployment in science, for example). I will not pursue this last kind of project, but there is certainly scope for utilising the ideas presented in this thesis in this direction, should the reader be dissatisfied with the goal I have aimed at here: namely a description of the actual concept we currently have, elucidation of the pressures that caused it to develop in this particular form, and a description and explanation of how we operate with it in everyday life.

### 1.3. The State of Nature; Informants and Sources of Information

In this thesis I will thus be pursuing what might be called a genealogical method: looking at the development of the concept of knowledge and the everyday conditions under which it arose. To this end, a useful tool in the naturalist’s arsenal is what we might call a ‘state of nature’ approach, which may be familiar from its employment in political philosophy, and is perhaps best described in an epistemological context by Bernard Williams in the opening chapters of ‘Truth and Truthfulness’ (Williams, 2002), or by Edward Craig in his book. This approach aims to develop what Craig has elsewhere called a ‘practical explication of the concept of knowledge’ (Craig, 1986/7).

In order to account for the universality of the concept of knowledge – Craig notes that ‘There seems to be no known language in which sentences using ‘know’ do not find a comfortable and colloquial equivalent’ – we begin with imagining a primitive situation: we consider the needs of some individuals living together in a (possibly quite primitive) society. They will be in possession of language, and some degree of cognitive sophistication, and have some interest in cooperating with one-another. One should not assume that individuals living in such primitive societies necessarily have what we might consider to be primitive needs; however, we shall start from consideration of needs of quite a general and basic kind, needs that inevitably emerge whenever human individuals live together in a society, and hence may be thought of as uncontroversial features of the human condition. If we find the explanatory power this approach to be lacking, we may then go on to consider needs of a more sophisticated kind. Given these basic needs, then, and some degree of reflective intelligence, we should expect that the concept of knowledge will emerge because of its usefulness in fulfilling them (the exact details of this ‘emergence’ will be the subject of the next section).

In order to meet the challenges of everyday life, each individual in our society must acquire information about the world, such the location of food, predators, competitors, and mating opportunities. Although false beliefs can occasionally be helpful, the best strategy for achieving their goals will be for them to aim at acquiring true beliefs about these variables. Such true beliefs can often be acquired directly through sensory experience and inference, and for most of our beliefs this is probably how we do indeed proceed. However, some individuals may be much better placed to provide input on a given question – because they possess special abilities, perhaps, or simply because they were conveniently located at the time of some event. The members of our society would therefore do well to share information with each other, thus distributing the epistemic burden in a cooperative drive for survival and prosperity.

To this end, certain concepts will be useful in determining which sources of information are reliable and which are not. Craig's proposal is that the purpose of the concept of knowledge is to 'flag approved informants' regarding the truth of some proposition  $p$ . On a technical note, Craig often talks as though our inquirer seeks to find out whether some proposition  $p$  is true or false, but the inquirer is not always in such a privileged position: I may be enquiring about the whereabouts of some individual whilst having no idea about where he is, and so in general my enquiry will not take on this disjunctive form. However, as long as we bear this in mind, we may adopt Craig's 'whether  $p$ ' disjunctive form of the question as it is far more convenient, and bears more similarity to the standard 'knows that  $p$ ' construction discussed in the literature.

I have said that the purpose of the concept of knowledge is to flag approved informants on a given question. There is a distinction between an informant and what is merely a source of information: a book may be a good source of information about many things, but we do not attribute to it the status of knowledge. How can we account for this fact from the 'state of nature' perspective? In other words, why not have a single locution for all sources of information, rather than a narrower concept that only applies to good informants? Craig's answer to this question is twofold.

The first reason is convenience: given that we speak the same language (I shall have more to say about channels of communication in the next chapter), once we have found a good informant we can find out whether  $p$  simply by asking an appropriate question, whereas a source of information may require expert or specialist knowledge to analyse. Someone who knows whether  $p$  is in some sense (other factors concurring) able to 'tell us whether  $p$ ': the concept of knowledge has a social function that would not be met by mere sources of information.

The second consideration, which rules out more convenient sources of information like books, is that a good informant will often act cooperatively by providing collateral information that is not explicitly requested: he or she may be able to ascertain or partly ascertain our purpose for wanting to know whether  $p$ , and are often empathetic towards our enquiry (this is especially true within a small primitive community in the state of nature). If I stop a man on a street in London and ask him where the nearest tube station is, in addition to answering he may also decide to tell me that the station is closed that day, or that there is no service on the Piccadilly line.

It is not merely that the entities to which we attribute belief must have agent characteristics, but rather that a good informant whether  $p$  – and hence one who knows  $p$ , *ex hypothesi* – must be able to convey to us whether  $p$  in a special way. To take Craig's

illustration, a man who is soaking wet may be a good source of information about whether or not it is raining outside simply because we can tell by looking at him. He may well also know that it is raining as well, but this does not follow straight away from the fact that he is both a good source of information and a fellow human being. We are using him as an informant only if he is able to *tell* us whether *p* – not simply if we infer it from facts about his physical appearance, say, which would be merely to make use of him as a source of information.

#### 1.4 Methodology

In the next sections I will be more explicit about the methodology we will be employing, and the form we should expect our theory to take. We have claimed that the role of the concept of knowledge – to point out approved informants – is somehow responsible for the conditions by which it is delineated. It is time now to be more explicit about the relationship between these two things. On page 10 of Craig's book we have:

'I shall not treat [the concept of knowledge's] development diachronically, and that is not just an omission: if what I shall say is along the right lines, the core of the concept of knowledge is an outcome of certain very general facts about the human situation; so general, indeed, that one cannot imagine their changing whilst anything we can still recognise as social life persists. Given those facts, and a modicum of self-conscious awareness, the concept will appear; and for the same reasons as caused it to appear it will then stay.'

It would be useful to try to give more detail about what is meant by 'an outcome of' here; likewise for 'the concept will appear'. Naturalism comes in degrees, and we may consider a variety of different proposals:

- To deduce the features of the concept it would rational to construct, given its purpose (the 'rational' framework)
- To deduce the features of the concept that is likely to develop organically as a practical strategy for ascription, given the conceptual niche available (the 'strategic' framework)
- To describe the features empirically found in real-life knowledge ascriptions, bearing our hypothesis about the role the concept plays in mind (the 'descriptive' framework)

In the first kind of project, we would try to make precise what is meant by a good informant, and then imagine constructing a concept perfectly suited to the role of flagging such individuals as part of a general epistemic strategy for acquiring true beliefs (what the concept of knowledge *should* be like). Should an individual possess all the properties this concept includes, then he or she must be a good informant by virtue of this fact. However, even in this level of analysis we must also take into account the fact that to be suited to this purpose the concept must be applicable by one situated individual to another (or, as we shall discuss, to oneself), as we shall see this will require the informant to be detectable (in the next chapter I shall discuss apparently counter-intuitive instances of people who we would like to say are knowers without giving any indication of this being the case). We are not rationally identifying particular individuals from a global

perspective, but rather imagining rationally constructing a concept for a certain purpose (I will elaborate on the importance of this point in section 1.6, below).

In the second kind of project, which is in many respects very similar, we consider the kinds of practical advantages we can gain from finding a good informant, and deduce the concept that would best help us to maximise them. The processes that cause this conceptual development may include evolution by natural selection: we have said that belief is a part of knowledge, and it is likely that biology has something to say about our belief-forming mechanisms. However, we are not limited to this framework alone; for example, those who find conceptual schema which lead them to material success of influence may use the power thereby gained to coerce others into following their usage – none of this resting on the idea that concepts are innate. At this level of description, the concept will (ideally) be the outcome of predictions from various branches of science, which deal with all the real-world forces that influence our conceptual practices.

Whereas at the rational level the concept would be tailored to ensure the individual in question was labelled a knower only if they were a good informant, at this level such a restriction is not in place: we want the concept to capture good informants *in general*, but it is likely to be strategically advantageous to allow some individuals who are in fact *not* good informants to fall under the scope of the concept in exchange for satisfaction of considerations like simplicity, ease of applicability, and ensuring that the majority of those who are good informants are labelled as such.

The third kind of analysis embarks upon the sociological project mentioned and criticized in section 1.2: to investigate how the concept of knowledge is employed in the everyday world. Two options suggest themselves, but each has their drawbacks. The first is to perform some kind of Gallup pole to catalogue intuitions regarding proper usage of the term ‘knowledge’. However, the answers we receive will likely not accurately reflect their conceptual practises, but rather some idealized version of it (as Lewis’s contextualism nicely captures, the conditions governing people’s ascriptions of knowledge often become more stringent at the mere *mention* of Cartesian demons). The second would be an analysis of large samples of recorded dialogue (spoken and written) to see how the term is used: but this kind of analysis cannot easily distinguish genuine usage from that which is merely metaphorical ascription and not to be taken literally, or ‘loose use’; that is to say merely a convenient approximation. These methodological problems will not particularly concern us, however: our advances in this direction will be restricted to a few platitudes about everyday usage that should prove intuitive and uncontroversial (such as that the term is frequently applied in real life).

Rather than committing to just one of these three styles of account, then, we may try to bear each of them in mind, carefully balancing our analysis of the core of the concept of knowledge between explaining the empirical facts of usage and deducing how we should expect them to develop or be chosen, given the various pressures to which humans living in a society are inevitably subjected. Clearly the sociological facts of usage are after not enough to go by; given that knowledge is factive it is an uncontroversial fact that knowledge is often ascribed incorrectly (some writers, such as Unger, whose views we shall consider in chapter 3 below, have claimed that almost all knowledge ascriptions are false). But conceptual schemas are not something that are wholly rationally constructed; nor are they the outcome of blind evolutionary and developmental processes, but more somewhere between the two. As mentioned in section 1, if the actual facts of usage differ

wildly from the concept we predict to evolve organically, perhaps we should take this to suggest that it is likely that these predictions are mistaken, and should at the very least question our assumptions or method to see where the discrepancy may have arisen.

With any luck we may find that some features are present (or not present) whichever approach we adopt, and by noting that this is the case – that all three approaches are in agreement – we stand on firmer ground in asserting these features to be part of the core of the concept.

One last point: we are concerned with knowledge in general, and not with any particular kind of knowledge – perceptual, mathematical, etc.: although usually  $p$  will be some fact about the immediate world around us, we have not specified in any more detail what it will be about. So if our hypothesis is correct, the features we delineate will be neutral as regards this also.

### 1.5. Form of the Theory

Before we proceed with exploring Craig's hypothesis that knowers are good informants, it will be helpful to say a few words about what kind of analysis we are looking for. As I have mentioned, the usual form in which an analysis is given is that of a list of individually necessary and jointly sufficient conditions, such that meeting them with respect to  $p$  and knowing whether  $p$  are mutually entailing. However, it also seems that whatever set of conditions one might give, one of two things happens. Either there is the possibility of 'Gettierization' – finding situations where the conditions are all met, but intuitively knowledge ascription should be withheld; or else the analysis becomes so demanding that is scarcely met at all in everyday life, if ever.

Does the approach of looking for features that characterize a concept suitable for flagging good informants help with this problem? Not directly, no: consideration of the features involved in being a good informant is unlikely to yield anything so precise as a set of necessary and conditions. Take belief, for instance: in the next chapter I will argue that for the most part a good informant will need to have true beliefs on the subject in question (a proposition  $p$ , say). But despite this, we can imagine cases such as Colin Radford's 'French Canadian' Jean (Radford, 1966/67), who as a result of prior instruction that has now passed from conscious memory is able to accurately answer questions on British history, despite apparently having no beliefs on the subject. Thus, including belief as a necessary condition would exclude Jean, who seems as helpful an informant as one could wish for, whereas to leave it out would seem to imply that the notion of belief plays no role in the analysis of knowledge – and this is surely problematic in itself.

What Craig does suggest to remedy the situation is a change of track. If it is correct that knowledge is a phenomena delineated only with respect to a particular kind of conceptual practise, and not by the limits of some natural kind, then there is no reason to expect the most illuminating analysis to even take the form of a set of necessary and sufficient conditions at all. Rather than looking for such a set, then, what we will instead aim to do is to look for a description of the *prototypical case*; that is to say, to highlight those features which are usually or almost always present when the concept is ascribed, even though we might be able to imagine exceptions:

‘what may look like an attempt to state necessary conditions should rather be taken as part of the description of a prototypical case, a case from which speakers and their audiences will tolerate, in the right circumstances, varying degrees of deviation. How much deviation, and under what circumstances, ought to be related to the purpose behind the formation of the concept in question. The prototypical description enshrines the features that effect realisation of the purpose when things are going on as they nearly always do.’ (p15).

In describing the prototypical case, we are looking for a description of those features that are usually present when the purpose that we have hypothesized the concept of knowledge to serve is fulfilled, and rarely present otherwise. Despite being closely associated with the fulfilment of the purpose, and hence part of the core of our inquirer’s conceptual strategy (and thus knowledge itself), these features are not logically necessary or even sufficient for adequate fulfilment of the purpose. Viewing the role of the concept as primary thus helps to explain disagreements about the wealth of purported counterexamples to popular analyses of knowledge that philosophers are so fond of constructing and citing as decisive: we may feel of a particular case both that many of the usual features of knowledge are present, but also that the role the concept is supposed to play is for some reason left unfulfilled. Conversely, one might construct cases in which the usual features of the situation are *not* present (such as Radford’s example) but there is an intuitive resistant to the concept of knowledge being withheld: ‘if our hypothesis is on the right track, it is neither surprising that so many take belief to be essential to knowledge, nor that some deny it, nor that many people’s intuitions leave them in the lurch at this point’ (Craig, p.14). Everyday practise gives us no way to legislate over these examples, and there is no reason why it should.

On a related note, the problem with the usual method of trying to imagine counterexamples to a proposed analysis is that this approach often does little to shed light on *why* the analysis is supposed to fail. Moreover, ‘imagine’ is perhaps the operative term here: the counterexamples that philosophers often like to emphasize, if genuine, are sometimes so rare that their impact on everyday conceptual practise is limited. Recall that our inquirer is concerned with a practical strategy: oftentimes it will be costly in time and energy for our inquirer even to *detect* those cases where knowledge is present but the usual features are missing, so in practice they will simply be passed over and a more accessible or readily identifiable informant sought instead. Therefore, the inquirer’s strategy will aim at the prototypical case, and indeed ‘one might almost say that for practical purposes what the concept amounts to is the essential description of the prototypical case.’ However, Craig goes on: ‘the words ‘almost’ and ‘for practical purposes’ are not in that sentence for nothing, however. One thing a proper account of a conceptual practise ought to be able to explain is why *prima facie* counter-examples to the proposed definition have (at least *prima facie*) the feel of such.’ (p15/16) Our attitude towards such cases, where they concern us, then, should not be to try to legislate based on ‘intuition’, but to shed light on these contrary pressures and how they lead to disagreement.

## 1.6 The Situated Viewpoint

In order to further clarify the differences between the naturalistic approach we are pursuing and more standard ways of proceeding, it will be helpful to introduce a

distinction between what might be called ‘global’ and ‘situated’ perspectives (c.f. Pfeifer and Bongard, 2007). An analysis beginning from a global perspective starts with all of the facts of a given (often imaginary) situation to hand, and goes about deciding whether ascriptions of knowledge are appropriate to a given individual. These facts generally include the truth as to whether  $p$ , the individual’s relationship to this fact, and various other details about his or her epistemic state and embedding in the world – even if these are not accessible to any individual present in the scenario.

In contrast, when taking the situated view, we imagine a real individual situated in the world, having limited information about his or her environment (including as to the truth of  $p$ ), and deciding whether or not to ascribe knowledge to another individual, despite not having all the facts to hand about their beliefs, intentions and history; or on the first-person approach (although I shall argue in the next chapter that the third-person approach is in some sense primary), deciding whether they themselves know whether  $p$ .

In this thesis I am pursuing an analysis from the situated perspective. Our inquirer will be embedded in the real world, which entails certain restrictions: acquiring information requires time and energy, sensory equipment is subject to disturbances, there is constant subjection to time pressures (things happen even if they do nothing); they interact with the environment, and these actions (including the overt application of concepts) change this environment in various ways. Further to this, the situated approach implies an active concern with one’s own practical interests: unlike Descartes’ ‘Pure inquirer’ – to borrow a phrase from Bernard Williams (Williams, 1978) – the situated inquirer begins with epistemic needs (amongst others) he wishes to fulfil. He then adopts a practical strategy aimed at doing so, and part of this strategy will be the way in which he uses and applies concepts.

Given this framework, the strategic approach asks: ‘which conceptual strategies might we expect to develop (by whatever means) to enable individuals in our society to flag good informants?’ – and the rational approach asks ‘if our goal was to flag good informants, which conditions would deliberation on this aim lead us to insist were present for conceptual application?’ And of course, the descriptive framework always takes place under real-world conditions.

Again, it will be worthwhile to pause here in order to address a concern that might present itself when we take this perspective; one analogous to that which was raised about the value of pursuing a naturalistic approach in general: why should we care about how knowledge looks from the situated viewpoint, instead of raising ourselves to a global one? Aren’t complications brought in by considering an individual’s particular situation at ground level mere prejudices, to be overcome by philosophical reflection?

The first point to make here is that it is the situated view that has been important in shaping the everyday conceptual application, and I have argued for the value of taking this conceptual application as primary. So if we are pursuing this naturalistic approach (and I have argued above that there is benefit in doing so) then this is the position that an effective enquiry at this level will have to work with. Moreover, the considerations that the situated viewpoint lead to cannot but bear influence on our intuitions about the proper application of the concept of knowledge – even when we try to divorce ourselves from it in our philosophical endeavours (for one thing, its influence will permeate our language).

A second consideration is that in any case our conceptual strategies do not fall apart once the inquirer gains a firmer grasp of the facts. For instance, oftentimes after using the

information obtained we will come to find out the truth about  $p$  after all, and supposing we had picked an informant who gave us the right answer then this would not then undermine their claims to knowledge in any way. More generally, we will find that the status of being a good informant is invariant with regards to increasing information on behalf of the inquirer: if some new facts serve to undermine his status as such later on, we will have been wrong to attribute the status of an approved informant in the first place. The account of knowledge aimed for is ultimately of such a kind that we will be able to say whether or not a given case qualifies from a global perspective (or at least whether it is part of the prototypical case). However, in developing this analysis it is considerations arising from the situated view that we will consider. Moreover, if we follow the classical tripartite analysis of knowledge, an analysis of belief and truth will come out the same from either perspective; as we shall see, it is only in motivating the third condition that this perspectival distinction will prove to be important.

We have seen how taking the situated rather than global perspective has lead us to construct our analysis from the position where we do not know whether  $p$  ourselves: to borrow another phrase from Bernard Williams, we are in the position of ‘inquirers, not examiners’. The situated view can also give an explanation as to why we should have counterfactual or modal concerns: knowledge is ascribed from a position of incomplete information, which marks out a range of ‘open possible worlds’ – that is, a range of worlds that are not excluded by our experiences hitherto. Because we cannot exclude the possibility that we are in any of these open possible worlds (that any of them are the actual world), our only hope of finding an informant with true beliefs whether  $p$  is to find one who has true beliefs across most or all of them (just how demanding we will be here remains to be seen, in chapters 3 and 4). From the global perspective we might think we only want our informant to be correct in the actual world, and in a sense this is true, but the problem with this is that it leads to no strategy on the part of our situated inquirer.

In general, as mentioned earlier, when we begin from the situated perspective we will also find that the features of the concept must be in some sense epistemically available in order to be worth bothering about for our inquirer. These points will be considered more carefully in subsequent chapters, and their importance for the debate we are addressing pursued.

## 2. The Concept of Knowledge

### 2.1 The Tripartite Analysis

As mentioned in the introduction, in this chapter I will describe the concept of knowledge as it emerges from consideration of our hypothesis, and then go on to categorize it according to various distinctions found in the literature. Lastly I will deal with some proposed counterexamples to the hypothesis.

The purpose of the concept of knowledge, we have said, is to point out individuals who are good informants whether  $p$ , in order to ourselves acquire true beliefs whether  $p$ . Moreover, an informant is distinguished from a mere source of information partly by a capacity to ‘tell us whether  $p$ ’. This given, we shall want to find someone who tells us the truth about  $p$ . Now, from this assumption Craig reasons as follows:

‘The informant, we may assume, will not in general tell him [the inquirer] the truth unless he (the informant) holds a true belief about it. (Cases of people who, whilst not holding true beliefs, insincerely give ‘information’ which is in fact true, are rare; and informants who do that regularly are as good as non-existent.)’ (p12)

Moreover, informants will only be of use to us if they express themselves with enough confidence to convince us: ‘if he is successfully to induce the belief that  $p$  in his audience, he had better believe it himself’ (p13).’ In the prototypical case, then, our informant will possess true beliefs whether  $p$ , and be sufficiently confident of this belief that once they tell us that  $p$  (or not- $p$ ), we will then ourselves come to believe  $p$  (or not- $p$ ). (As mentioned, we may agree with Radford that the belief condition is not strictly necessary, but it is nevertheless so common that we may take it to be part of the ‘core’ of the concept of knowledge: see chapter 1, section 6.)

We might initially think that nothing further than true belief is needed by our practical inquirer, and in a sense this is right: all that they really need in a good informant is for his or her views to be true and for them to have the confidence to express them convincingly. However, recall that the purpose of our naturalistic enquiry is to find the practical conditions under which knowledge is ascribed (or better, *correctly* ascribed). As I have mentioned, most of the time we will not know the truth about  $p$  ourselves, or else we would not bother to look for an informant. What we are really after is someone that believes  $p$  if  $p$  is true, and not- $p$  if  $p$  is false (the similarity with Nozick’s tracking account of knowledge will be pursued in the next section). We must be able to *distinguish* informants with true beliefs about  $p$  from our situated perspective where there are all sorts of other things about the world we simply do not know, and be able to do this without ourselves having prior knowledge as to whether or not  $p$  is true.

In the background of other conditions that also correlate well with being a good informant (such as being to some degree rational), the fact that the informant has a belief whether  $p$ , or is confident about whether  $p$ , is a good indication that what he or she will say on the subject is accurate, and hence that we should come to believe them too. Indeed, this is part of the motivation behind our emphasis of the belief condition: we do not look for confident informants because we find confidence attractive, but rather

because, other things concurring, it is a good indicator of correctness (it is thus only of indirect importance). However, confidence is not the only indicator we use, and not even one we are always willing to trust: there are other properties a potential informant may possess that indicate their having true beliefs (or false beliefs) whether  $p$ , and some of these may be more reliable. At this point, it seems that any such property will be sufficient for the purposes of the inquirer; call this the ‘minimal analysis’ – that knowledge is true belief together with some detectable property indicating that this true belief condition is met (whether we really need this ‘detectability’ constraint will be addressed in section 2.6, below).

One question that we will be concerned with throughout the remainder of this thesis is whether to express this property (that indicates true beliefs whether  $p$ ) in modal or probabilistic terms (Craig often conflates these two notions; for example, on p86). Are we after an informant that is correct across a certain range of possible worlds – perhaps across all worlds that, for all we know, do not obtain? Or are we after someone who is (either objectively or epistemically) very *likely* to be correct? These two notions do not coincide, but there is at least the semblance of affinity between them: for instance, usually someone who is known to be correct in almost every possible world is also likely to be correct in the actual world. This question will be answered in chapter 4, but for now we shall remain non-committal on this issue and stick to Craig’s formulation – that what our inquirer is after is:

‘[S]omeone who believes the truth about  $p$ , and has some detectable property X, possession of which correlates well with being right about  $p$ , that can guide him in his choice of informant.’

In particular, the property must be detectable to persons lacking direct access to the truth about  $p$ . Moreover, any such property will do, so long as the agent conveys the information with the ‘informant’ characteristics outlined above: this property X will not be specified in any more detail, as there are as many such properties as things we could want to know about. If we want to know how to get to Piccadilly Circus, the – quite readily detectable – property of being a London Taxi driver will suffice. Just what is involved in this ‘correlating well’ clause will occupy us a great deal throughout the thesis: indeed, in a sense this question contains the essence of the debate under consideration. One last point before we move on: the connection between this property X and being right about  $p$  must be reliable or law-like as well as graspable, and not just a *de facto* connection that holds in this particular instance (otherwise the inquirer would not in general be able to discern that the informant was in fact suitable for his purposes).

## 2.2 Comparison with other Projects

In chapters 3 and 4 of his book, Craig makes a comparison between the minimal analysis as presented in the previous section, and four other attempts to analyse the concept of knowledge: the Nozick/Dretske counterfactual tracking account, Goldman’s causal theory, reliabilism, and the classical ‘justified true belief’ (JTB) analysis. Although I will not be able to pursue this in as much detail, a few brief remarks will be helpful at this juncture.

The reader will recall that in the Nozickean tracking account, the third condition is decomposed into two subconditions, namely *sensitivity*: that in possible worlds that are close (i.e. worlds very similar in *p*-relevant respects) to the nearest possible world to the actual world in which not-*p* obtains the subject does not believe *p*, and *adherence*: that the subject's beliefs are the same in all close possible worlds where *p* has the same truth value.

Craig has two comments to make on this approach. Firstly, according to our practical explication as outlined so far, Nozick selects the wrong range of possible worlds: there is no reason for our inquirer to care about worlds he has ruled out as being non-actual. Although we shall see in section 2.6 that things become more complicated when we consider that our inquirer will in general not operate alone but from within a linguistic community, on the simple analysis it appears that all an inquirer should care about is what the potential informant believes across the range of possible worlds that is open to him, the inquirer. Under the strategic framework, an open possible world is best understood as one that has not been excluded as a candidate for being the actual world by that subject's experience; under the rational framework, it is a world that they do not know is not the actual world. Consideration of some such set of worlds will automatically bring the counterfactual property in its train: recall that in general our inquirer will not know whether *p* obtains or not, and hence this set of open possible worlds in which the subject is required to be correct will contain both worlds where *p* is both true and false.

The second problem with Nozick's account is that the truth of these counterfactuals is not epistemically accessible in the appropriate way: our inquirer has no convenient faculty for judging whether these conditions hold. So at best satisfying the counterfactuals will be something that correlates well with the property we are looking for: it may be present nearly all of the time that the property *X* is present, but it cannot be the property itself (again, whether this detectability requirement is really necessary will be considered in more detail in section 2.6 below).

Next, Craig discusses Goldman's causal theory of *a posteriori* knowledge, which analyses knowledge whether *p* as true belief caused by the fact that *p*, with the proviso that the subject be able to reconstruct the salient links in the causal chain. There is a good deal of similarity to Nozick's project here, as a causal connection is prone to bringing counterfactual conditions in its train. And again, it appears that according to our methodology this formulation also misses the mark for the same reason: this property will not in general be epistemically accessible to the informant. In order to judge whether a belief is causally connected to the fact that it is a belief in, we will usually need to know that the fact actually obtains; and of course our inquirer is rarely in such a position. Granted, there are cases in which this is not the case: we may believe the causal connection holds because we saw him looking in the right place at the right time, for instance (p14). But in these cases it is some other feature of the informant we use to ascertain that the causal connection holds – here simply that he was looking in the right place – and it is this other, more basic feature that we are really looking for.

Let us now look at the reliabilist formulation, which analyses knowledge as true belief acquired by a reliable method. It is true that for a nearly all of our beliefs there is something like a method behind their acquisition, and if this method is reliable then the subject must be a good informant, given their beliefs are correct. But as far as looking for a property that a subject is in possession of which correlates well with being right about

*p*, the property of having come across *p* by a reliable method is only one such indicator; we have already mentioned another we are likely to sometimes use in some cases, namely being confident about one's beliefs. Other kinds of properties, such as being a London taxi-driver, do the job just as well – and being a London taxi-driver is not a property that can be equated with coming about with certain beliefs about the local geography of the city, even though it is true that all competent taxi-driver will have used some such method.

According to the state of nature perspective, then, our inquirer will not so much be concerned with the method by which the beliefs are acquired, as by whether there is something about the informant that indicates that their opinion on the matter in question is going to be reliable. What they want is for the detectable property X to be reliably connected to being correct whether *p*. This given, counterfactual and causal conditions may creep in (but in a non-standard way) in ensuring this condition holds: if being reliable (in a law-like way, recall) requires causality (Craig thinks not), then there will be necessarily be a causal element involved; if it involves the satisfaction of counterfactuals (this time Craig thinks it does) then these will have to be met too. But again, as these conditions are indirect consequences of other parts of the analysis, we may stick with the minimal formulation as it currently stands. Likewise, to consider the JTB analysis, having a good reason for a belief is a reliable indicator that it is correct, but not the only such correlate.

According to Craig, all four of these approaches are good approximations to the minimal analysis, which seem plausible because in so many cases the detectable property X happens to be something that makes them come out true (especially given what we have said about the importance of the prototypical case). As Craig puts it:

‘What makes them seem plausible is not the concept of knowledge, but certain very general beliefs which we all hold. These are, in particular, beliefs about the extent to which the world is a system of causally inter-related states, more specifically beliefs about the extent to which belief-states are themselves the end-product of a causal process; the belief that for nearly all human beliefs, there is such a thing as the method by which they were acquired; and the fact that human beings are usually conscious of certain stages of the processes by which they arrive at beliefs.’ (p34)

### 2.3 Globalism versus Localism

In this section we will discuss the first of three distinctions that have been drawn in the literature in order to categorize various epistemological approaches, and will be instrumental in developing the arguments against infallibilism found in the next section. The globalism/localism distinction is usually discussed in terms of reliabilism: recall that the reliabilist approach to knowledge claims that knowledge is true belief acquired by a reliable method, whence the question arises of whether in order to count as leading to knowledge whether *p* the method must be reliable for determining the truth of a range of other *p*-like propositions. For our purposes, this issue turns on whether to count as knowing whether *p* it is necessary that S is also a reliable informant on a range of other *p*-like questions (presumably only in a localist sense, to prevent explosion), and if this is the case, which range is the appropriate one.

As Craig points out (p.54), McGinn (1984) has objected to Nozick's account of knowledge on the grounds that it only refers to the specific proposition *p*, and includes no

clause that quantifies over a range of other  $p$ -like propositions. However, Craig also points out that there are *prima facie* counterexamples to the claim that any such global reliability is required: an infant might know its own name despite not knowing the name of anyone else. The later example may be taken to suggest that the descriptive framework indicates globalism is false: but it is not clear that statements expressing the names of other individuals are the appropriate class of  $p$ -like propositions in this instance; perhaps what the infant also needs to know is a further range of facts about *himself* to count as really knowing her own name. Likewise, I might know the answer to  $17^{20}$  because I have worked it out on paper, whilst not knowing the answer to any similar-looking calculation such as  $17^{19}$  or  $17^{21}$ ; but maybe the range of things I need to know about here are not other calculations but more general mathematical principles, such as the mechanical procedure of calculation, the fact that when properly implemented it leads to true results, and so on.

However, if we admit (as many authors do) what must be consequence of our hypothesis, that knowledge can be transferred from an authoritative source (perhaps insisting that the individual learning the information be aware of their standing), then there will be a large number of examples where globalism does not obtain; when the authoritative source on the subject (a professor specialising in that field, perhaps, or a quizmaster on a television show backed by a team of researchers) tells another individual some isolated fact  $p$  for which there is no range of related  $p$ -like propositions the learning subject is aware of. Because knowledge attributions of this sort are common, specific conditions relating to globalism are not part of the core of the concept according to the descriptive framework.

Let us, then, consider things from the rational framework. Supposing we are looking for a good informant whether  $p$ : would it be rational for us also to insist that they are also likely to answer correctly on a range of other  $p$ -like questions? On the face of it, it seems they are not: we are only interested in learning whether  $p$ , and to include extra conditions about other  $p$ -like issues would only serve to exclude individuals perfectly well-suited to our purposes. However, recall that one of the two key reasons for only ascribing knowledge to good informants rather than otherwise convenient sources of information such as books was that a good informant was often able to spontaneously provide collateral information relevant to our concerns whether  $p$ : if we want this to be the case, perhaps some version of globalism will need to apply as a consequence.

If there were such a condition motivated for this reason, it would either have to be either very broad (the subject needs to also know about every possible contingency that could be relevant to our concerns with  $p$ ; which is surely too strong, given the belief requirement), or the range of  $p$ -like propositions specified will be highly contextually determined, and dependent not just upon the semantic content of  $p$  and our concerns with its truth, but also the specific details of the circumstances surrounding our relation to it (all the things we in particular, and at that particular time, might need to know collateral information about). However, from the strategic framework there is an alternative way in which we can see to it that this concern with the possibility of gaining collateral information is satisfied, so that again there is no reason to think that globalism is correct.

Our inquirer, we have said, seeks to find out the answer to some question  $p$ , and hence also seeks a good informant that amongst other things has some (detectable) property  $X$  that correlates well with being right about  $p$ , which has the property of being law-like, in

order to support inference to new cases. Now, in the majority of cases, possession of the same property X will also correlate well with being right about a range of *p*-like issues: if Sam is a good informant about whether the cat is on the mat because he is looking directly at it, he will also be a good person to ask about the shape and size of the mat too.

Combined with this observation, we may note that it is usually easier to find a subject with a more general correlative property than a specific one: in order to know the date of a particular battle we look for a subject with a firm grasp of the relevant period of history, and not simply one that knows this one particular fact. Thus strategically we will do better in general to *look* for individuals that also know about a range of *p*-like propositions, if only as a consequence of the fact that we are looking for informants in possession of more general correlative properties than is strictly necessary. This range of *p*-like propositions will depend not on the semantic content of *p* itself but rather the means by which we as inquirers come to know that a subject knows whether *p*.

We thus have seen that there is no reason to expect our situated inquirer to directly care about the informant's reliability on anything other than the question at hand; reliability in these other cases is a bonus that comes about from the way we identify good informants whether *p*, rather than something that is part of the core of the concept itself. Hence as the globalistic properties only come about as a consequence of other elements of our analysis – that the individual must possess some property indicating that they are likely to be correct whether *p* – we may leave a description of it out of our constructed concept, and opt for localism about knowledge instead.

#### 2.4 Internalism Versus Externalism

The internalist/externalist distinction turns on whether in order to count as knowing whether *p*, the subject needs to themselves have some grasp of the fact that they fulfil the third condition for knowledge. Hence for us the relevant questions are whether to include any clause to the effect that S must grasp that they possess the correlative property X, and the related question of whether they need to themselves grasp this connection between possessing X and being a good informant whether *p*.

In the next section I shall argue that a third-person (as opposed to first-person) account of knowledge is the firmer starting point. Now, from the third-person set-up – i.e. consideration of the case where an inquirer must decide whether an external informant should be ascribed the status of knowledge – it initially seems that there is no reason to think internalism about knowledge is correct. All that is required of a potential informant is that they possess the appropriate property; whether or not they realize this to be the case is irrelevant.

However, recall that we are also required to detect such informants, and moreover the reasons behind distinguishing good informants and mere good sources of information were related to the convenience of being directly told whether *p*. Two points are relevant in this regard. The first is that it might be true that sufficiently confident belief whether *p* is usually (in almost every instance) accompanied by a grasp of the facts that justify this confidence (i.e. that the subject possesses the property X). This given, the internalist condition would have something like the status of the belief condition: whilst not logically necessary for being a good informant – it is certainly possible to be very

confident without a grasp of the reason underpinning this confidence – it may nonetheless be part of the prototypical case, especially where confidence is an appropriate attitude to take. Against this line of thought, Craig counters:

‘But be that as it may, we have still found one point on which the supporter of the belief-condition can call which has no analogue for the internalist: the lack of belief seriously disrupts our willingness to think of the subject as an informant rather than an evidential source of information. In part, at least, this happens because we are then reluctant to regard him as properly *telling* us whatever it is that he says. (p.63)’

Moreover, the inquirer has no direct concern with the informant’s grasp of the correlative property in regards to its supporting an attitude of confidence; all that matters is that the informant has the required property and believes confidently enough to attract our attention (i.e., the two conditions are separable, and the latter need not be based on the former). Hence these reasons are not enough to suggest that our inquirer’s conceptual strategy will have a built-in internalist condition.

A second point to note, and one that I think bears more weight, relates again to the convenience of a good informant. Recall that what our inquirer wants is someone that can properly ‘tell them whether *p*’, as opposed to merely being a source of evidence in regards to their question. But if this emphasis on convenience is correct, perhaps we should also insist that they are able to tell us that they possess the correlative property X, rather than being expected to infer this from detectable facts about them. And telling us that they possess X (in the sense we are interested in) requires them to appreciate this fact – hence the internalist’s condition follows.

In response, we might ask why someone who accepted this line of thought should stop here: an informant can certainly express his belief that he possesses the correlative property X, but we need a reason to think *this* utterance is reliable. Again, either we can look for some property X\* that correlates well with being reliable about *this* sort of thing, or we can insist that the inquirer reliably tells us that he or she possess such a property. But then the same point can be raised again, and so on. Where on this scale should we stop?

Again, I think the answer is that we should stop before the first iteration, insisting only upon belief whether *p*. To argue this, I will appeal to the three frameworks of analysis discussed previously. Beginning with the descriptive framework, oftentimes we do ascribe knowledge when the internalist condition is not met. This is the outcome of two more general facts: that subjects possessing such a property do not always consider whether or not this is the case, and further that we do not in general have access to the facts about whether a subject has considered the issue or not. Hence in some proportion of the cases where we ascribe knowledge the subject will possess some such property X, but will not meet the internalist condition.

From the rational stance, it seems clear that all a practical inquirer really needs from a good informant in order to find out whether *p* is that they possess some such property X; grasping this to be the case is not necessary, and it would not be rational to exclude informants simply because of the fact that they do not. Hence the rational stance also favours externalism.

Lastly, let us consider things from the strategic point of view. If the subject is aware that they possess the property X, then surely this will be useful to us in identifying that they

do possess it: for example, the question of ‘how do you know?’ may then enable us to determine whether or not they are a good informants based on their response. So would our inquirer do well with a strategy aimed at just these kinds of good informants?

The answer, I think, is no. The subject expressing to us that they possess some such property X is only one way we come to identify that this is the case; it is not a dominant prototypical feature that occurs in almost every instance (other ways we might come to see that they possess such a property include awareness of their social or professional position, our own knowledge of their whereabouts at a particular time, direct perception, the testimony of others, and so on). Hence, a strategy aimed exclusively at informants satisfying the internalist condition would only allow an inquirer to make use of a small subset of potentially helpful informants.

Lastly, we can now also answer the second question raised at the start of this section: given that we do not require the subject to be aware that they possess the correlative property X, there would be no reason for us to require their awareness of the fact that just this property correlates well with true beliefs whether *p* either. Again, it may *useful* (given they know they possess X) to suppose they are aware of the connection; it might make them more ready (or more able) to identify themselves as good informants whether *p*. But it will not be necessary, as in general there are other ways we can come to know about this connection.

## 2.5 Third Person and First Person Perspectives

Philosophers often distinguish between first person and third person accounts of knowledge. In the first-person case, we take as the paradigmatic case of knowledge ascription the situation where we apply the concept to ourselves, and express (to others or ourselves) that we know some fact *p* (or not-*p*) to be the case. Conversely, by taking a third-person analysis we have in mind the case where, seeing some external subject, we decide to ascribe to him knowledge whether *p*, thus endorsing him as a good informant on the subject (either for our own clarity or, as we shall see in the next section, for the benefit of others). Which of these two approaches our methodology favours will depend on the relative influence that judging others versus judging ourselves as good informants have had in shaping the associated conceptual practices which we have claimed are the origins of the meaning of the term ‘knowledge’. In emphasising the distinction in this section I am not claiming that different concepts will ensue when adopting one style of explanation over the other, but simply looking for a clear starting point that is as firm as possible.

So far we have essentially had implicitly in mind a third person investigation, but the ‘state of nature’ methodology is not necessarily limited to this approach: indeed, oftentimes it is me that is best placed to evaluate and subsequently put myself forth as a good informant. However, I will argue in this section that we should take the third-person scenario as our starting point.

Craig suggests one potential line of justification for this emphasis on third person ascription – that first person judgments require more quite sophisticated reflective abilities: ‘To find oneself in possession of a belief on the question whether *p* pre-empting inquiry; to take a self-conscious look at one’s own apparatus with the doubt in mind that it may have delivered a falsehood calls for a considerable degree of sophistication.’ (p11)

He claims that we should pursue the third person approach, then, because the less cognitive capacity a phenomena requires imputing to our individuals, the less controversial is the claim that it is an inevitable feature of life in any human society. By being conservative in the abilities necessary to find oneself in our starting position, we thus broaden our enquiry: even if we should later find a reason to move to the first-person view, until such reason arises it is ‘just good method’ to keep our approach as general as possible.

However, although the introspective approach may require sophisticated abilities some of the time, it seems that it needn’t always: for instance, suppose I see what I initially believe to be a tiger in the distance, but subsequently take into account how the foggy weather means that this belief is unreliable. I have evaluated my status as an informant, and hence the applicability of knowledge, without doing anything particularly sophisticated at all – certainly nothing that could not be done in the state of nature. Moreover, it is well known that infants below a certain age have no ‘theory of mind’ – they are unable to fully understand external subjects having false beliefs, whilst understanding the possibility of their own views having been false. If this is correct then Craig’s argument is not particularly convincing. Let us, then, pursue the question of which style of account we should pursue using the frameworks from chapter 1. If the concept of knowledge following from the starting point we will adopt disagrees markedly with actual usage this will become apparent later on, so let us move on to the developmental approach.

Now, although it is in fact not clear that it requires more complex cognitive abilities, it seems correct that third-person application does enjoy numerical advantage: it is simply far more common for us to become involved in judging whether another individual is a good informant than to find ourselves making the same judgment about ourselves. Even granted that most of our beliefs are formed directly, as opposed to from the testimony of others, we do not necessarily need to employ the *concept* of knowledge in order to acquire these beliefs, and although we may be an appropriate subject for knowledge ascription there is no reason why we should have to step back and reflectively assess our status as informants in more than a very small proportion of these cases. Moreover, it is the inquirer that initiates the dialogue regarding the particular question at hand, and has easier access to facts about the circumstances surrounding his enquiry (I shall have more to say about these later).

This given, and assuming the two diverge at all (I think in fact they probably do not), then a concept designed to meet the demands of third-person ascription will be in general more useful for the inquirer to possess, and so this is the concept we may expect to develop under the strategic framework. Moreover, the same result follows if we consider which concept it is rational to construct. If our hypothesis is correct, then the only reason we could have for constructing a concept that labels ourselves as good informants would be to recommend ourselves as such to others. But then supposing there is some divergence in how the concept will develop between the two styles of account, it will be of no use for the semantics of ‘knowledge’ to be suited only for a first-person scenario in these divergent cases, because the point of the first person version of the concept is to be able to present oneself as a good informant to others. Presumably everyone will be willing to assert that their own beliefs are correct; others will in general seek independent

grounds to believe us. In short, the importance of this usage of the concept is derivative from the third-person case.

For these reasons, the best conceptual strategy will aim at the third-person scenario as its prototypical case, and we may expect the meaning of the term 'know' to follow. Hence in the following sections and chapters we shall take as the prototypical case of knowledge-ascription that of attributing knowledge to an external individual, as this is the 'firmer theoretical starting-point'.

One last point before we move on: in 'Knowledge and its Limits', Timothy Williamson has put forth the influential view that 'knowledge is the norm of assertion' (Williamson, 2000). In other words, one should only assert what one knows (or – following Stanley, 2005, whose views we shall consider in more depth in chapter 4, one should only act on what one knows). This leads to an alternative hypothesis with more of a first-person character, which might likewise be pursued along Craigean lines: that the purpose of the concept of knowledge lies in evaluating whether we are in a position to make certain assertions. But if what I have said about the third person approach being more fundamental is correct, exploration of the current hypothesis is best completed ahead of this project, although it may certainly have a role in tidying up any phenomena the current project leaves unexplained. Moreover, the two hypotheses are not contradictory to one-another: there is clearly a close relationship between being a good informant whether  $p$  and being justified in asserting whether  $p$ .

## 2.6 Objectivisation

Up until this point, we have been working with a rather simplified picture of information transfer in our primitive society: we have imagined an isolated inquirer seeking potential informants amongst the other individuals living in his or her society, without any input or help from others. However, life in a society is not quite like this: for example, our ascriptions (or withholdings) of knowledge are often performed publicly, which in turn affects how others act. In the ideal case, there is also a mutual desire for cooperation – or at least for successful communication: if we want to understand what others are saying, we must to some extent take their situation into account. In this section I will consider how complications such as these are likely to affect the concept of knowledge.

One might think that if the purpose of knowledge is to flag approved informants, then we should expect the truth conditions relating to knowledge attributions to be highly context-dependent with regards to the circumstances surrounding enquiry. These circumstances include our purposes for wanting to know, the degree of information we already possess that might be relevant, our attitude to risk, and so on. On p.85 Craig explains how an inquirer requires an informant that fulfils the following properties (in addition to the three already mentioned, apart from I4, which relates to it):

**I1: They should be accessible to me here and now**

**I2: They should be recognisable by me as someone likely to be right about  $p$**

**I3: Channels of communication between him and me should be open**

**I4: They should be as likely to be right about  $p$  as my concerns require (Craig, p85. My labelling/reordering)**

Clearly, the fulfilment of each of these conditions is relative to details of the particular situation in which our inquirer finds himself or herself at the time in question. However, life in our primitive community is inherently social: I may find myself in a position to recommend a good informant to James, say, because James is himself required to recommend a good informant to Susan – and we may then suppose that I have no knowledge of the circumstances surrounding Susan’s enquiry. Moreover, many knowledge ascriptions take place with no particular purpose in mind; merely forming part of the backdrop of common societal knowledge about who is *in general* a good informant about what (clearly we have an inherent curiosity of such things; hence the popularity of quiz shows). Therefore, in our analysis we may subject each of these conditions to what Craig calls ‘objectivisation’, whereby we abstract away from details of the particular circumstances surrounding one individual’s enquiry to arrive at a more substantial and intersubjective (but still anthropomorphic) concept. Our individuals living in a linguistic community will thus be able to convey information about someone’s status as a potential informant *in general* about whether  $p$  concisely and unambiguously, using the single locution ‘know’.

In this section I will explore how each of the first three conditions becomes de-relativised: condition I4 will be seen to be of special relevance to the fallibilism versus infallibilism debate (although the idea behind it is often expressed using a *modal*, rather than probabilistic vocabulary) with which we are concerned, and will be discussed in chapter 4.

Let us first consider condition I1. As well as using the concept to flag informants who are accessible to me here and now, it will be useful to identify informants that may be accessible to me in the near future, or if I am willing to move a short distance. Furthermore, an informant may not be accessible to me, but she is accessible to Fred, and I can find out what she thinks from Fred. But because pretty much any spatial or temporal location a potential informant might find himself or herself is in theory reachable by another, it will in fact be more useful for members of our society to apply the concept to all informants that otherwise qualify, regardless of whether they are accessible here and now. In this case objectivisation proceeds ‘all the way’ – in the fully objectivised concept there is no mention of the location of the informant at all.

Let us now consider condition I2. Again, another individual in the community might recommend an informant to me that, whilst actually suiting my purpose, I would not have been able to spot myself. We may likewise extend the scope of our concept to those good informants that are (in principle) detectable by some human means. Should we extend it any further, or drop the ‘detectability’ condition altogether? Whilst noting the verificationist line which ‘warns us, on pain of insignificance, not to let this ‘could be’ clause (in ‘could in principle be detectable’) become too strong’, Craig considers the possibility of there being correlative properties lying without the scope of actual or even possible human capabilities to detect (Craig, p89).

Let us approach this question using the three frameworks. Firstly, the descriptive one: as humans we cannot systemically designate individuals with properties not detectable by human abilities (even as they currently stand), because flagging them as such would be

counted as detection of the property. Secondly, there could be no practical advantage to a strategy aimed at including them, because again *ex hypothesi* we cannot detect them. Lastly, we cannot rationally decide to look for informants satisfying such a condition, given that our search would be futile. Hence unlike I1, if our hypothesis is correct then the detectability condition will leave its mark on the concept: good informants must in principle be detectable (we will see the importance of this remark in the next chapter). This means that the correlative property X must be one which is in principle detectable by possible human means, and the concept as we currently operate with it will pertain to actual, current human means. Moreover, given the situated nature of the enquiry and the derived emphasis on the third-person scenario, we may stand by the claim from section 2.1 that this property X must be (in principle) detectable by an external subject who themselves lack access to the truth about whether *p*.

Lastly, let us consider condition I3: the channels of communication between me and another individual may be shut, but they can easily be reopened should I manage to find a third person that can act as an interpreter or mediator. Hence I would be unwise to exclude a person as a potential informant on this basis alone, and as before we may extend the concept to any individual with whom some human specialist may be able to communicate. But again, objectivisation does not proceed all the way: recall the first reason for only ascribing knowledge to informants rather than sources of information – they were convenient to communicate with. A similar analysis would follow from our three frameworks.

## 2.7 Proposed Counterexamples

We are now in a position to offer a response to some *prima facie* counterexamples to the claim that the purpose of the concept of knowledge is to flag good informants. Craig speaks of two problematic cases in particular: the gangster Luigi, who knows where Mario's body is, but will not say; and those who know certain facts but will not be believed because they lack credibility: for instance, the boy who cried wolf, or Matilda from Hilare Belloc's 'Cautionary Tales'. Similarly, a trained spy taken prisoner by a hostile nation may be said to know certain government secrets – he has a detectable property that indicates he is likely to be right about these sorts of things – whilst being no use as an informant, because he will not tell.

It seems that in these cases the individuals in question are not good informants, but nonetheless are intuitively said to know; and so there is a divergence between the intuitive extension of 'know' and that which is demarcated by exploring the consequences of our hypothesis about the role of the concept. However, objectivisation suggests a response to these cases. It seems the problem lies in their failure to meet I3 – the channels of communication are shut – but this condition features in a far weaker form in the objectivised concept. Hence, Luigi may not tell me where the body is, but perhaps he will tell Carlo, and then there is a chance I can find out from him. Likewise, the majority of the townsfolk may not believe Matilda, but perhaps someone with a more intimate knowledge of her psychology would be able to tell she was not lying on this occasion (she will also likely be believed by anyone with *less* of a connection to her that was unaware of her past mistruths). Similarly with our spy; we can always imagine contrary pressures which change the playing field in such a way to alter his priorities:

perhaps he is given an incentive he cannot refuse; perhaps he changes hands and is willing to tell his new captors. Thus we can see why intuitively it would not do to rule out these individuals as potentially helpful in our enquiries on principle; they may be the only ones who are able to help.

Let us now consider some counterexamples to the condition I2. Note that the individuals in our previous examples met the unobjectivized version of this condition: the problem wasn't that they were not recognized as having true beliefs on the subject in question, but rather with the possibility that they would not convey those beliefs sincerely. Consider firstly the 'secret murderer': a man who has murdered his wife, managed to make it look like an accident, and taken it upon himself not to ever tell anyone or reveal any sign of this being the case for as long as he lives. It seems that there is no outward sign that he is a good informant about her true cause of death, and yet intuitively one thinks he must know that he has killed her. But can both of these things really be true? He might not have any property correlative with beliefs about the actual truth that is detectable by the people he encounters in his daily life; but perhaps if we avail ourselves of a lie detector it might be. Failing that, perhaps in theory we could use complicated neural imaging technology to identify neural patterns that indicate that his real beliefs.

If he if he is immune to all such tests that humans are in capable of devising, and so shows no external or internal marks to indicate his beliefs on the matter are actually the correct ones (any part of his behaviour that indicates that he knows might be accidentally observed), then it seems more appropriate to surmise that he in fact does *not* know: perhaps he has convinced himself that it was not him that killed his wife, in which case he fails the belief requirement. In support of this, we find the view in Dennett (Dennett, 1989) as well as Fodor (with a precursor in Ryle, 1949) that one cannot really ascertain what one believes via introspection, but only through a systematic analysis of one's behaviour. For instance, in Ryle we find the example of an ice skater who claims to believe it is safe to venture toward the centre of a frozen lake, but his actual beliefs on the matter are exposed by his tendency to ensure himself and his family stay away from it when skating (Ryle, 1949). And if beliefs must be supported by corresponding external behaviours, then they are always in some sense detectable.

Lastly, consider those exceptional human individuals who cannot communicate at all, due to complete paralysis, say, but may still be said to know various details about their lives (a brain-scan reveals normal activity). If rare counterexamples like this are indeed genuine, they are unlikely to bear influence on our conceptual practices (c.f. chapter one: the 'prototypical case'), especially given their inability to participate in normal social life.

A more general form of these concerns is what Craig calls the 'cart before the horse' objection, which points out that a good informant is someone that knows whether  $p$  only if circumstances concur. Therefore knowledge is the more primitive concept, and being a good informant should be analyzed subsequently as knowledge plus some other conditions governing this concurrence. Now, although the considerations in this section have brought the two notions closer together, we still need not claim that an identity holds between being a knower and a good informant for our hypothesis about the role of the concept to be correct, or for the Craigean methodology to be a valid and useful one. Craig illustrates nicely this with another example: the purpose of a chair is that it is a convenient object for individual humans to sit upon. Some chairs (for example, the one atop a lit bonfire) do not serve this purpose, despite not losing their claims to chair-hood;

but nonetheless an enquiry into the existence of chairs would do well to take this hypothesis about the purpose of chairs as its starting point. Likewise for the epistemologist:

‘By envisaging such a story we do several things which the proponent of the ‘cart before the horse’ objection does not. If we say: ‘Knowledge first, and then comes the capacity to inform’, we leave it obscure why the concept of knowledge should ever have arisen, rather as the comparable objection about the concept of a chair would leave that concept, and even more so the chairs themselves, unaccountable brute-fact bits of the mental and physical habitats.’ (Craig, p. 97)

What objectivisation provides is an explanation for how it can be both true to say that the concept of knowledge has its origins in the role it plays in everyday life, and further that conditions relating to the concurrence of circumstances that are necessary for fulfilling this role appear only in a weakened form; or, in the case of I1, do not appear at all.

# 3. Some Varieties of Epistemic Infallibilism

## 3.1 Unger and Scepticism

In this chapter, I will describe a number of formulations of epistemic infallibilism, and aim to show that each one is too strong to be necessary for knowledge ascriptions. In the next chapter I will complete the argument against infallibilism in general by motivating a positive fallibilist position about the concept of knowledge.

In his 1971 paper 'A defense of skepticism', Peter Unger attempts to demonstrate (that it is reasonable to believe) the truth of the thesis of scepticism – that all anyone knows is at most a very small number of propositions – by showing that there is a condition for knowledge (absolute certainty) that is necessary for proper ascription but almost never met in real life.

He marks a distinction between terms he calls absolute – such as 'flat' – and terms that are relative, such as 'bumpy'. Something is only properly called 'flat' if it is to no degree curved or bumpy, so that 'flat' and 'absolutely flat' are synonymous, whereas bumpiness is a property that may occur in varying degrees. Although absolute terms like 'flat' appear to admit of qualification – it is natural for one to describe a football pitch as 'very flat' – these utterances are best paraphrased with constructions like 'very close to being flat', so the appearance of their having a relative character disappears. Moreover, no such paraphrase is available for the relative terms: to say something is very bumpy is by no means to say it is 'close to being [absolutely] bumpy'. He has a variety of other examples to illustrate the distinction:

I think that the first term of each of the following pairs is a relative term while the second is an absolute one: "wet" and "dry," "crooked" and "straight," "important" and "crucial," "incomplete" and "complete," "useful" and "useless," and so on. I think that both "empty" and "full" are absolute terms, while "good" and "bad," "rich" and "poor," and "happy" and "unhappy" are all relative terms. (p207)

For practical purposes we often extend usage of absolute terms to other things; even though a purported vacuum might not literally be devoid of matter, it is useful to refer to it as though it is, if it is empty enough for all practical purposes. However, this is dismissed as mere pragmatics: philosophical analysis should aim to capture proper usage in the ideal case. It is thus not just knowledge, but whole classes of absolute terms that are systematically used incorrectly by being applied when they should not be.

It is clear from the pairs of examples above that oftentimes an absolute term can be defined in terms of (the total absence of the property corresponding to) a relative term. Unger gives such an analysis of the term 'certainty': he claims it is the same thing as the complete absence of doubt (doubting being a 'relative' phenomena, as one can be more or less doubtful). His argument for scepticism is then that knowledge requires one to be certain of the proposition that one knows, and it is very rarely (if ever) the case that one is

absolutely certain with a complete absence of doubt about any given proposition; even one as pedestrian as ‘ $45+56 = 101$ ’.

It may seem at first that Unger’s analysis of ‘certainty’ in terms of absence of doubt is quite trivially flawed: after all, it is surely intuitively possible to lack completely any feeling of doubt over certain propositions that one has not come close to ever considering, but it would be a gross misuse of language to say one was certain of them (for one thing, we might have a similar attitude of unconsidered indifference to the *negation* of such a proposition). However, this point does not ultimately hurt Unger’s position: recall that another necessary condition for knowledge we have derived is that the subject *believes* the proposition in question, which excludes the counterexample cases where one is simply indifferent.

Unger’s argument for the necessity of certainty as a condition for knowledge ascription is linguistic: he claims that sentences of the form ‘I know  $p$ , but I am not certain of it’ are infelicitous. Let us take this claim to be true; we must then decide whether to understand certainty in Unger’s absolute sense, or in a wider, more variable sense that indicates a high degree of confidence. The situation seems analogous to that with knowledge: we have an intensional intuition (let’s say) regarding what it means for a belief to be certain – that it requires complete and total absence of doubt – and also some extensional intuitions: in relation to many propositions we have the *feeling* of certainty as a distinctly recognisable mental phenomenon. Certainty in this non-absolute sense is somewhat opposed to and undermined by doubt, but in the same way as happiness and sadness are antagonistic: the two mental phenomena are mostly likely based on distinct physiological processes, which while mutually inhibiting do not require a complete absence of their opposite to be properly instantiated.

As with knowledge, we may try to resolve this dilemma with the method of hypothesizing a role for the concept. And at least within this context, we can see that the purpose of the concept of certainty must be to have available to us a way of indicating a high degree of confidence in our beliefs. Given this is correct, then the broader sense of ‘certainty’ suggested by our extensional intuitions lead us to an alternative explanation of the linguistic data: the function of self-knowledge ascriptions, we have said, is to advocate oneself as a good informant whether  $p$ . We have also noted that in order to be of any use as an informant, we will in general need to be sufficiently confident of our beliefs to induce others to follow them. Now, if we understand ‘certainty’ in the broader sense, as merely indicating a high degree of confidence in our beliefs, then we can see why these sentences are infelicitous: by admitting we are not certain (in this sense) we are suggesting (in contradiction with the first clause of our utterance) that we are in fact *not* good informants. The problem with the sentence, then, is not so much to do with its logical truth conditions as that it expresses two contradictory messages.

Fortunately I do not think we will have to decide which analysis of ‘certainty’ is correct; for even supposing that the absolute analysis is a coherent concept, that doubt as a psychical phenomena can be precisely quantified and that there are psychological states  $S$  could be in which correspond to its complete absence, is it surely quite unreasonable to suppose that one has conscious access to this being the case (as opposed to, say, an *almost* complete lack of the absence of doubt) and hence the influence of these kinds of states on our actual conceptual practises regarding the term ‘certainty’ will be minimal. Our intuitive reaction to statements of the form ‘I know  $p$ , but I am not certain of it’ does

therefore not count as evidence that knowledge requires this absolute kind of certainty, as this intuitive reaction must be a product of certainty in the wider, less precise sense, as this is the only kind we ever knowingly encounter.

In the remainder of this chapter we will no longer be directly concerned with certainty in the psychological sense of this attitude towards one's beliefs, but rather with the manner in which our beliefs come to be, or are guaranteed to be, correct.

### 3.2 The Exclusion Principle

In this section I will discuss what we will call the 'Exclusion Principle', which is often formulated as follows: if we are to lay claim to knowing that some proposition  $p$  is true, we must be able to exclude every possibility that is logically incompatible with  $p$ . This principle has been advocated by David Lewis<sup>2</sup>, amongst others:

If you are a contented fallibilist, I implore you to be honest, be naive, hear it afresh. 'He knows, yet he has not eliminated all possibilities of error.' Even if you've numbed your ears, doesn't this overt, explicit fallibilism still sound wrong?' (Lewis, 1996)

The formulation just given introduces an element of internalism: the subject must himself or herself be able to exclude these possibilities (whatever *this* might mean), the active tone of which implies they have some awareness that this is the case. Given we have argued against the necessity of internalist conditions in the last chapter, we will consider instead an externalist version of this idea<sup>3</sup>: that S's epistemic situation (to be made more precise shortly) must be incompatible with any possibility inconsistent with  $p$  – that is to say, the truth of  $p$  is logically implied by his or her epistemic situation. For example, S may satisfy an evidential condition such that it is impossible for him or her to satisfy that condition whilst  $p$  is false. More generally, in any possible world where S's epistemic situation remains a constant,  $p$  is logically required to be true.

This idea seems intuitive, but will require clarification. As mentioned, we must make more precise what is meant by 'epistemic situation': whether this is taken to be demarcated by something subjective, such that being in the same epistemic situation as S is equivalent to being some entity with the sum total of S's impressions and introspectively accessible thoughts; or, less generally, having S's subjective mental state in that precise instantiation (such that if physicalism is true this would require their neural state to be identical); or a function of S's mental state, S's specific embodiment in the world, and the relationship between them.

We will consider what happens if we consider the third possibility later on in this chapter (following Williamson and Dretske – see section 3.6); for the moment we will assume that one's epistemic situation can be analyzed internally and then later fused together with the factive condition that the subject's beliefs about  $p$  are actually correct.

---

<sup>2</sup> As we shall see in the next chapter, Lewis goes on to qualify this to 'relevant' possibilities – the scope of which are contextually determined.

<sup>3</sup> For convenience I will still use the term 'exclusion principle' to denote this condition: the possibilities being excluded (in the passive sense) by the subject's epistemic situation, not the subject themselves.

In the next section I will argue that the weaker formulation of the exclusion principle – that  $p$  must logically follow from our psychological state in that precise instantiation – is false, and hence the stronger formulation also. In order to do this I will argue that oftentimes knowledge that  $p$ , where  $p$  is some fact about the world around us such as ‘there are more people living in London than in Cambridge’, does not require one to exclude the possibility of the familiar ‘Brains in Vats’ scenario; for it is consistent with our internal neurological state that this is the case: there is a possible world in which a brain precisely identical to ours in its current state is envatted and fed input giving rise to the appearance of it having all the experiences we hope to be having now.

Before going on to give this argument, some other comments will be appropriate at this juncture. Firstly, we may note that the principle as given includes no counterfactual condition; we have said nothing about what  $S$  would believe, given that  $p$  were in fact false. Although this condition is only supposed to be necessary for knowledge, and not sufficient, it would be possible for us to avoid the need for unnecessary further conditions by translating it into the terms of the account of knowledge presented in this thesis: that we may only correctly ascribe knowledge to individuals in possession of some detectable property  $X$  such that possession of that property logically excludes the possibility of their being incorrect *whether*  $p$ .

More generally, we have here a further direction in which we can vary the possible worlds we are interested in: so far we have been keeping  $S$ ’s epistemic state fixed and varying possible worlds around it, but we could also range over possible worlds where  $S$  is in a different epistemic state. Presumably  $S$ ’s existence is contingent, and no-one would require that  $S$ ’s beliefs about  $p$  were correct in worlds where  $p$  did not exist, but we could insist, for example, that  $S$  was such that any state  $S$  could be in logically implies correct beliefs *whether*  $p$ , ranging over all possible world where personal identity held between  $S$  and the subjects in those worlds. Thus not only must  $S$  be in a state logically implying  $p$ ;  $S$  must be such that they could not *but* be in such a state with regards to  $p$ . Of course, if – as I shall argue – the weaker version as characterised here is false, this stronger version must be too – but this kind of counterfactual condition might still play a role in the final constructed concept.

I will close this section by noting that the line of thought just explored also suggests an argument against the exclusion principle as formulated above. I will not pursue this argument very far, because I wish to avoid giving a justification for its key assumption, as this would take us too far afield into other areas of philosophy. However, I shall briefly outline it, both to anticipate the actual argument given in the next section, and to consolidate the ideas given so far in this one. Supposing it is true that considerations of possible worlds where  $S$  does not exist cannot be directly relevant to the question of whether  $S$  is a good informant *whether*  $p$ , and hence to whether  $S$  knows  $p$ . If we put this thought together with the (well-advocated, but controversial) claim that facts about  $S$ ’s physical embodiment in the world are required to be held constant for personal identity to obtain, then we need not consider worlds involving BIV, where there is an alternative embodiment for the central subject of impressions (which will *not* be  $S$ ). But then it is possible that  $S$  both knows that  $p$  is true whilst their belief that  $p$  is incorrect in some of these worlds; hence the exclusion principle as formulated is false.

### 3.3 Against the Exclusion Principle

Using the three frameworks of enquiry outlined in the last chapter, I will here claim that correct knowledge ascription is possible whilst the exclusion condition (that  $S$ 's epistemic situation must logically entail  $p$  is true) is not met because it is often not rational to demand that it is met, the concept will not develop to feature it, and the empirical facts of usage suggest we do not in general require it either.

The last of these three claims is easily seen to be true: as Unger correctly notes, if the exclusion condition was part of the concept of knowledge then we would know very little, if anything at all. But this is clearly not how the term is used – ascriptions of knowledge about the ‘imminent world’ (see the next paragraph for a definition of this term) are common – and so actual usage does not demand we take note of the exclusion condition. So let us move on to the question of whether we should expect an organically evolved concept of knowledge to include this exclusion principle as one of its conditions.

I use the term ‘imminent world’ to refer to the realm of direct experience a subject has (following e.g. Valberg, 2007): hence for us the imminent world is currently the everyday world around us which most of our beliefs are about, and which contains London, Berlin and Australia. For a dreaming person, the imminent world is the dream world; hence a subject’s imminent world need not actually exist. If BIV obtained, then, the imminent world would not exist and hence almost all of our beliefs would be false. Consequently, we would know at most very little.

Now, explanation at the strategic level aims to use various branches of science to shed light on conceptual practises as the outcome of organic processes. But within the special sciences, as well as evolutionary theory, these are processes that happen here in the imminent world around us (hence if this world does not exist these theories are false). For instance, the theory of evolution by natural selection is a historical theory specifically about the occurrence of processes in the imminent world. This given, the practical advantage our strategies aim at, explained within these frameworks, will make no reference to the possibility of its non-existence: there could, for example, be no evolutionary survival value in the ability to exclude the possibility of a BIV scenario, because survival value in evolutionary theory is something that is explicitly connected with the imminent world around us. Hence analysis at this level suggests that in order to know  $p$ , where  $p$  is some everyday proposition about the imminent world, it is not necessary for the subject to exclude the possibility that they are a BIV, and therefore that the exclusion principle is false.

In make this claim I am not endorsing this and saying we *should* only care about the imminent world: my project is to describe the concept of knowledge as it currently stands, and the special sciences and evolutionary theory are important frameworks that we can to use in this endeavour.

Lastly, let us consider whether it is *rational* for us to build in the exclusion principle to our concept of knowledge, given the purpose that we want it to achieve. Let us again (with Craig) take BIV as a representative sceptical possibility, and assume that  $p$  is some proposition about the imminent world that we have some practical concern in finding out about. In the last chapter I argued that we should take as our prototypical ascription scenario the case of deciding whether an external subject  $S$  is a good informant whether  $p$  (i.e., a third-person analysis). In order for this question to make sense to us, we must

believe that this external subject S exists – which we will generally only come to believe if we believe that the external world through which we appear to have access to them exists. Moreover, supposing again that *p* is some question about the imminent world; we would ourselves only maintain a practical interest whether *p* so far as we were convinced that this imminent world did in fact exist. Hence the question of whether to ascribe knowledge to S, when put in this manner, is set firmly within a context where we are ourselves convinced the external world exists, and with the subject S positioned within it.

But this being the case, it can be of little interest to us whether or not S can exclude the possibility that they are a BIV, because in order to make sense of their testimony, we must already be steadfastly convinced they are not. It may often be helpful to us if S *believes* that the external world exists: indeed, by being actively cooperative in their role as informant they are in a sense acting as if they believe it to. But there is no reason we should demand more than true belief on their behalf when inquirer whether *p* (where *p* is some practical question). It will therefore be of no concern to an inquirer that a potential informant does not possess a convincing proof of the existence of the external world, in the style of Descartes' famous meditations. Given our other interests, then, we are not rationally required to insist upon S meeting the exclusion condition: indeed, given the difficulty of actually producing a successful refutation of BIV, it would be positively detrimental to our practical concerns were we to do so.

The most that the infallibilist can demand, then, is that there is a 'real world' certainty of the subject's being correct: in Lewisian terms, when *p* is some proposition about the imminent world we have a practical concern with, the range of possible worlds across which we require true belief from the subject does not include any of the worlds where BIV (and many of the sceptic's other favourite possibilities) obtain. In Craig's words:

[e]ven when we do set the likelihood [of the subject's being correct] at 1 that does not mean the type of absolute certainty, invulnerability to literally any theoretical possibility, which the sceptic characteristically demands and finds us unable to provide. Nothing about our reactions to lotteries or the danger of treachery need push us any further than the demand that the actual or 'real' chance of being wrong should be zero. Hence I suggest, the attractiveness of the idea that someone knows whether *p* when something about him is connected as a matter of natural law with his holding the right belief about *p*.

We shall take up this idea of the subject being correct as a matter of natural law in the coming sections; for now we merely note that the fact that knowledge ascriptions about a wide variety of propositions are possible without the exclusion principle being met is enough to assure us that it is not part of the prototypical case, and hence we may leave it out of our analysis of the concept of knowledge.

### 3.4 Knowledge and Accidents

In his 1968 paper 'An Analysis of Factual Knowledge', Peter Unger advocates the following analysis:

**'For any sentential value of *p*, (at a time *t*) a man knows that *p* if and only if (at *t*) it is not at all accidental that the man is right about its being the case that *p*.'** Unger, 1968, p.158

Although initially understanding this notion in a way which makes knowledge whether  $p$  a rather easy condition to meet, Unger later suggested that the problem with his previous views was that he interpreted the notion of ‘not at all accidental’ too liberally: ‘My main error, then, was not that of giving too vague or liberal a defining condition, but rather that of too liberally interpreting a condition which is in fact strict.’ (Unger, 1971, p12.) As we shall see in the next section, interpreting it strictly we get another kind of infallibilism, and one which it will be fruitful to explore (by considering whether it is necessary for knowledge). But how should this ‘accidental’ clause be understood, and what reasons for thinking that it is important will our practical explication suggest? Let us address the former question first.

Unger illustrates a sense in which it is *not* to be understood: ‘In my analysis of human factual knowledge, a complete absence of the accidental is claimed, not regarding the existence or abilities of the man who knows, but only as regards a certain relation concerning the man and the fact.’ (p.159) Hence one may come to know by accident, of a fact that it itself an accident, whilst at the same time it not being to any degree accidental that one’s beliefs about this fact were right. In terms of our analysis, we can express this by saying that the property X, which indicates reliability whether  $p$ , should do so in a way that is to no degree accidental. Hence if  $p$  is a contingent matter – and in general it will be contingent, else we should not be enquiring – the property X should guarantee correctness about  $p$  whether  $p$  comes out true or not, and again counterfactual conditions are brought into the analysis.

At this point, we can (following Craig, section 6) motivate Unger’s condition in a plausible way by connecting it with another feature from the literature: it seems that whatever analysis of the third condition is given, one can usually construct (Gettier-style) counterexamples such that all three conditions are fulfilled, but the subject’s beliefs are only true ‘by accident’, and this undermines our willingness to attribute knowledge to the subject. But why is it that the third condition (which, remember, we motivated via a need to identify informants with true beliefs) need to ensure that the other conditions are met in a specifically non-accidental way?

We have said in the last chapter that the property X should be law-like, in order to enable an inquirer to tell *in advance* that there is a high chance of a potential informant’s beliefs being correct. However, Gettier-style examples are constructed in such a way that some unusual feature is present (or usual feature is absent) which makes the usually reliable correlative property X unreliable in that particular instance, whilst at the same time it actually does lead the subject to having correct beliefs, but in an entirely accidental way. An example will make this clearer.

Supposing I am standing outside the door of the UCL seminar room, and through the door I can hear the distinctive voice of my favourite professor. I thus come to believe (correctly) that the professor is inside the seminar room: however, unbeknownst to me, the individuals in the room are actually listening to a recorded debate of some philosophical issue, and it just so happens that the voice issuing from the device’s speakers was that of my favourite professor.

In this example, the connection between hearing the professor’s voice through the door and being right about whether they were inside is usually reliable – this may have been the first time such an audio device was present in the seminar room. However, in this

instance it was entirely an accident that my beliefs were correct (we may suppose he was supposed to have caught a flight earlier that day, but it was cancelled due to some other accident, so he was able to attend class after all). Why should our practical inquirer exclude this kind of informant? Supposing someone had asked me whether the professor had been in the room, and later found out that the information he or she had received was correct, as well as the other details surrounding how I came to believe this: what reason would he or she have to be dissatisfied? Craig suggests:

‘It produces that retrospective feeling of having run a risk, of having done something that one would not have done had one just been a little better informed at the time, rather like finding that the person who has just driven you 50 miles down a busy motorway without incident hasn’t passed the driving test.’

In one sense, in using me as an informant the inquirer would have been unlucky that the correlation did not have its usual connections to the truth about  $p$  in this instance. However, they would have been in another sense terribly lucky; it was actually not wise to use me as an informant on this particular instance, and if they had done so they should count themselves very lucky not to end up with a false belief.

To a situated inquirer, perhaps observing me listening by the door, things would appear normal, and hence they would be inclined to use me as a good informant (and thus end up with a true belief). However, when we describe the case from the global perspective we are left with the feeling that this would be an unwise move, despite it leading to a true belief for the inquirer, because we can see the sense in which it was only accidental that my beliefs were correct. It is thus understandable that we should intuitively want to withhold knowledge ascription if doing so is tantamount to recommending me as a good informant. For we should not want to be in the position of having to rely on such a person in these particular (abnormal) circumstances ourselves.

### 3.5 Against Unger’s Non-accidentalism

We can thus see that when we come to analyse the concept of knowledge we will find reasons for not wanting our beliefs to be true merely by accident. But does this mean that we should require the property  $X$  to be connected to the truth about  $p$  in a way that is *to no degree* accidental? One (sufficiently non-liberal) reading of this is to require that something about the subject is connected to the truth about  $p$  as a matter of natural law – indeed, if this condition fails, then there must be some contingent fact unconnected with  $X$  or  $p$  that is enforcing the connection in this case. It may seem rather a stretch (if not an outright abuse of language) to count any independent contingent feature as an ‘accident’ – but there is no absolutist condition that falls short of this; just varying degrees of accidentalness. We can imagine a series of intermediate cases: perhaps the class intended to listen to the debate on the speaker, but it had blown a fuse, so the professor had to read out the transcript instead. The subject being correct by virtue of natural law lies at the extreme end of this scale, where his beliefs are to no degree accidental – and anything short of this would not be infallibilist in the required sense.

If this is the condition, that given the subject’s internal physical state it follows that upon any possible embedding of them in a Universe subject to the same physical laws

their beliefs would come out true, then I think it is too strong to be necessary for knowledge. For one thing, our individuals in the state of nature have no faculty for distinguish such cases; as Craig points out (p102/3), the brain's belief-forming mechanisms are highly complex. Rather, given a connection that is in general reliable, even though there might be conceivable situations in which this connection could systematically fail, and circumstances could conspire to make the individual in question wrong whether  $p$ , it is still possible for someone to be a good informant (simply) by virtue of the corresponding property. As Craig puts it:

'Gettier cases draw attention to the fact that even very good reasons indeed can let you down; but if that be allowed to tell against Smith's credentials then it tells against everybody else's. Of course, it doesn't do either. The correlation between having excellent reasons and being right is still what it always was: fallible, but virtually unfailing.'

For example, consider the case where, a week earlier, I had heard the actual voice of my professor through the door and no tape recorder was present. Now, this time I think that I *would* have known he was in the room, even though it was actually possible (given my epistemic state, understood in an internalist sense) that he was not present and the tape-recorder scenario obtained. If Unger's position entails exclusion of this kind of case, then this absolutist kind of non-accidentalism seems wrong, because the kind of factor that motivated us to withdraw our consent to ascription – i.e. the unusual features of the situation that make the correlation accidental in this instance – are absent, even though given the property and our epistemic situation it is (in some sense) an 'accident' that these features were absent. And as noted, there is no obvious absolutist-style cut-off point between the usual case of hearing through the door and the deviant case with the recorder.

When determining whether a subject knows whether  $p$ , we usually hold fixed some other background features of the situation that are not specified by a description of the physical laws. As we shall see in the next section, this practise receives support from Dretske, but it can also be motivated from the state of nature account: given there is some feature of the world that is public knowledge, ubiquitous but still in some sense contingent, we would do well to adopt a strategy aimed at knowledge ascriptions taking place with the truth of this feature assumed as a background condition. These background assumptions must be widely appreciated by everyone in the community, including the subject themselves, of course; they cannot be of a specialized or technical nature, known only to the inquirer and a handful of others, or else we cannot properly consider the informant as 'telling us whether  $p$ '. To take a modern example of a *de facto*, in some sense accidental, but nevertheless reliable connection, GPS systems may exploit some coincidental and arbitrary feature of the cosmos, and yet an individual in possession of one may be connected to the truth of his location as a matter of natural law (as well as veridical sense perception of the display).

In the next chapter we will support these intuitions and somewhat fragmentary ideas about Unger's condition being too strong with a positive argument for a more lenient condition, exploring how the non-accidental character of the connection between  $X$  and the truth of  $p$  can be captured in a fallibilist way will generally include (for example) putative cases of knowledge such as hearing through a door in the usual case. The

remainder of this chapter will focus on the ‘conclusive reasons’ formulation of Fred Dretske.

### 3.6 Conclusive Reasons; The Broadness of the Mental

We have seen in section 3 above that entailment from one’s epistemic situation, conceived as the sum of one’s subjective experiences, is too strong a condition to be necessary for knowledge. Now, picking up a hint from the last section, we will consider the position advocated by Fred Dretske. Dretske has claimed that knowledge ascriptions imply that the knower possesses ‘conclusive reasons’ – evidence which, in a sense to be clarified (together with the relevant sense of ‘possession’), do entail (either logically or via the laws of nature) that the proposition in question be true. However, ‘evidence’ and other epistemic notions are understood in a broader sense than we have been treating them hitherto, and the extension of the concept of knowledge thus demarcated is far greater than the meagre allowance typically granted by the sceptic (indeed, we might see Dretske’s position as a kind of response to scepticism).

In this section, I will sketch out the Dretske line, something similar to which has also been advocated by thinkers such as McDowell. Firstly, it will help to introduce some terminology, following Williamson’s paper ‘The Broadness of the Mental’ (Williamson, 1998). Let us begin with the mildly physicalist assumption that, given a particular agent *S*, their total internal (i.e. lying within their physical boundary) and external states (lying outside this) determine the total state of the world. Suppose further that we can treat these two kinds of states separately (if the physical laws mean that this cannot be done, the point follows anyway). A ‘case’ is a temporal slice of a possible world, centred around a distinguished subject. A ‘condition’ is something that may fail to obtain in a given case, and is specified by a ‘that’ clause – such as ‘the subject is happy’ (which obtains in a given case if and only if the centred subject of that case is happy).

A case  $\alpha$  is ‘internally like’ a case  $\beta$  if and only if their respective subjects have identical internal physical states, and a condition is ‘narrow’ if and only if it obtains in  $\alpha$  whenever  $\alpha$  is internally like some other state  $\beta$  where it obtains (it only depends on the internal). Conditions not meeting this requirement are ‘broad’. Many conditions relating to mental phenomena are broad: for example, seeing Naples is broad because someone internally like someone who was seeing Naples is not therefore seeing Naples themselves; they could be subject to an illusion, for example (Williamson, 1998, p390). Conditions relating to external states in the same way that narrow clauses relate to internal states are called ‘environmental’. A condition is ‘composite’ if and only if it may be expressed as the conjunction of a narrow condition and an environmental condition, and ‘prime’ otherwise.

So far it may sound as if we have been treating knowledge as if it were a composite condition that can be decomposed into the factive condition plus some manifest properties of the observer. But consider the following kind of argument, which shows that knowledge is in fact often governed by prime conditions.

To show that a condition is prime, it will do to provide cases  $\alpha$  and  $\beta$  such that the condition obtains in both, but does not obtain in a case  $\gamma$  which is specified to be internally like  $\alpha$  and externally like  $\beta$ .

Let  $\alpha$  be a case in which one knows by testimony that an election was rigged. We are told this by Smith, who we trust. Brown also tells us the election was rigged, but he is not trustworthy and we do not trust him. Conversely, let  $\beta$  be a case in which both Smith and Brown tell us, but this time it is Brown that is trusted and trustworthy, whereas Smith is distrusted and untrustworthy. Now consider the case  $\gamma$ , which is internally like  $\alpha$  and externally like  $\beta$ . In this case no informant is both trusted and trustworthy, hence we do not know that the election was rigged. Hence the condition that we know about the election is prime, and it is clear that examples with this structure can be constructed to say the same thing about pretty much any kind of knowledge that relates to the external world (Williams presents similar arguments for other mental predicates such as ‘seeing’ and ‘believing’).

In his paper ‘conclusive reasons’, Fred Dretske claims that a necessary and sufficient condition for knowledge is possession of ‘conclusive reasons’. Reasons R are conclusive for  $p$  if and only if given R,  $\neg \diamond \neg p$  (where  $\diamond$  expresses some hitherto unclarified modal operator). Alternatively, we can formulate this as ‘given R,  $\diamond (R \ \& \ \neg p)$ ’. This means that condition 2. below is entailed by condition 1.:

**1. S knows that P and he knows this on the basis (simply) of R**

**2. R would not be the case unless P were the case.**

A person *has conclusive reasons*, R, for believing P if and only if:

**(A) R is a conclusive reason for P (i.e. (2) is true),**

**(B) S believes, without doubt, reservation, or question, that P is the case and he believes this on the basis of R,**

**(C) (i) S knows that R is the case or**

**(ii) R is some experiential state of S (about which it may not make sense to suppose that S knows that R is the case; at least it no longer makes much sense to ask *how* he knows). (Dretske, 1971, p12/13)**

The definition is saved from circularity ((C)(i) contains a reference to knowledge) because recursive application will eventually reduce knowledge ascriptions to conditions of the type (C)(ii).

Now, the Dretske line is immune from the arguments for infallibilism presented above on several counts. Firstly, the modal operator  $\diamond$  does not denote logical entailment of P from R, or even entailment via the laws of nature, but entailment (either logically or nomically) only relative to a particular set of circumstances. Dretske explains:

‘If someone remarks, mid-way through a poker hand, that if his neighbour had not folded (dropped from the game) he (the speaker) would have been dealt a royal flush ... He is not saying that his neighbour’s remaining in the game is, quite generally, sufficient for his receipt of a royal flush. Rather, he is saying that *in the particular circumstances which in fact prevailed on this occasion*, circumstances which include such things as card distribution, arrangement of players, etc., an occurrence of the first sort (neighbour remains in game) will invariably be followed by one of the second sort (he receipt of a royal flush).

Likewise, in general when we consider claims to knowledge we are to hold fixed those circumstances that are causally and logically independent of P and R, and see if it is possible for P to be true whilst S has reasons R (and there is some resemblance here to our interpretation of Unger's position).

Secondly, as mentioned above, the notion of evidence is understood more broadly (in Williamson's sense) than we have been considering hitherto. Consider an individual sitting outside in their garden, and seeing a robin perched on the fence. Given that they see it, is it possible that the robin is not actually there? As the sceptic is fond of pointing out, there are possible worlds in which there are individuals having the same subjective experiences as this person, whilst the robin is not really there at all (they are being somehow systematically misled about this sort of thing). However, if the conditions governing notions of seeing are characteristically prime – and Williamson argues convincingly that they are – then the individual in question is not *seeing* the robin in these cases either (this is essentially the disjunctivist line). Likewise, neurological twins are not necessarily in the same epistemic situations, because evidential relations are prime. For example, a BIV with an identical brain to yours is not in the same epistemic situation.

On this understand of epistemic notions, then, the sceptical challenge does not show that the evidence is inconclusive, because in the sceptical worlds where P is not true the subject does not possess R. The sceptical challenge only undermines the question of whether we *know that we know*: usually veridical experiences are all that is required for knowledge itself. Therefore, if most of our experiences are veridical, and we should like to hope that this is the case, entering into the knowledge relation is an everyday, commonplace occurrence. In the next section I will discuss how the Dretske position thus outlined ties together with Craig's project.

### 3.7 The Craigean Response

I have said that conditions imposed on knowledge ascriptions by Dretske allow the concept to have a much wider extension than that typically granted by the sceptic; indeed, perhaps even something close to the actual usage of the term 'knowledge' from the descriptive framework (which, on our approach, will have to play a large role in determining what counts as an adequate analysis). Because it does not define knowledge into being something exceptional and rarely attained, and leaves room for knowledge as a phenomenon that actually plays an important role in everyday human life from its most primitive forms, the Dretske position is not infallibilist in the sense that this thesis aims to refute. However, *prima facie* it seems that there are at least two places where Craig and Dretske's analysis diverge. The first is that our earlier insistence on the informant having a *detectable* property does not seem to fit harmoniously with the idea that epistemic notions such as evidence are prime, so that whether or not they are instantiated is indeterminate given only the internal state of the possessor (whereas one might *prima facie* think a detectable property should have to be determinate in this respect). A second, related difference is that we have said that this property should be detectable to individuals who do not themselves have access to the truth whether *p* – which seems to imply that it should be independent of the truth about *p* – whereas according to Dretske's

views, whether or not someone has ‘conclusive reasons’ depends on whether or not their experiences are veridical (and in order to ascertain this, given only a report of their subjective impressions, we might need to have access to the truth about  $p$  ourselves). The remainder of this chapter will aim to show that, although the two positions do diverge, the contrast is not as radical as may be thought at first sight.

As I have mentioned, the fact that the property  $X$  (which indicates being a reliable informant about  $p$ ) needs to be (at least in principle) detectable seems to imply that on the Craigean account the third condition for knowledge is narrow (in Williamson’s sense; see section 3.6, above). However, this inference would be overly hasty. Recall that our analysis is proceeding from the situated viewpoint, with informant and inquirer standing in some specific spatio-temporal relation to one another, rather than the inquirer reading, say, a printout detailing the potential informant’s internal physical constitution. The relevant sense of ‘detectable’ is therefore one that naturally brings in its train certain implications about the physical embodiment of the informant with relation to the inquirer. A typical example might proceed as follows:

Mrs. Jones, sitting in the living room, sees Mr. Jones staring out of the kitchen window. After a brief exchange in which it emerges he is staring at the moon, Mrs. Jones asks her husband if the moon is full that night. Now, the detectable property that suggests he is likely to be right about this question is not that he is internally constituted in a certain way, so as to give rise to certain subjective experiences, but rather, it is specifically that he is staring out of the window at the moon. This can be directly observed by Mrs. Jones, who can also see that he is not hooked up to any elaborate visual device designed to deceive him, and can thus infer that his experiences are veridical and (given that channels of communication are properly open – i.e., there is no question of deliberate deception) his reports accurate.

More generally, what will often occur in the prototypical case is that when the property suggesting an informant to be reliable is detectable, the means by which it is detected (prior to a knowledge ascription) will also be such as to suggest that the property is in fact coupled to the environment in such a way as to make the experiences of the subject (the potential informant) veridical. Moreover, in judging whether their beliefs are likely to be true, we only rely on evidence presented by an informant insofar as we believe it to represent veridical experience.

It is true that narrow conditions must feature as the basis for any strategy we ascribe to the inquirer – indeed, we can understand a strategy as a conditional response to how things seem; the strategy is in general only working effectively when these ‘seemings’ are veridical, but this condition does not feature as part of the strategy itself. However, this consideration applies only to the inquirer, whereas the informant – the prospective knower – will usually be an external person. This given, there is no reason to think conditions governing ascription of the detectable property  $X$  to be composite; indeed, if Dretske’s arguments for prime epistemic notions connecting better with action are valid, this is the opposite of what we should expect. Consideration of the state of nature thus provides good reasons to think the evidential notions are prime. And if we do understand both the property  $X$  itself and ‘detection’ as concepts that are characteristically prime, the divergence between the two accounts is greatly reduced.

Moreover, on this understanding of ‘detectability’, the second reason for which the two accounts may be thought to diverge is also softened: understanding the property  $X$  as something that is only instantiated if the experiences or reasons the subject has which

connect it to the truth about  $p$  are veridical, detection of the property X (which must be in some sense possible, remember) does suggest the informant is likely to be right about  $p$ , and this can be done from the inquirer's position of not knowing about  $p$  themselves. Detection thus only occurs if the property X is actually instantiated, which requires it actually having the appropriate connections with the surrounding environment. And an inquirer can see this to be the case without knowing whether  $p$  themselves.

Lastly, although it is not infallibilist in a way that would obviously conflict with the picture of knowledge as I have constructed it so far (as opposed to, say, a concept involving the exclusion principle), Dretske's position still has something of an absolute character: knowledge requires reasons that provide logical or empirical entailment (holding fixed circumstances independent of those reasons and of  $p$ ) of the truth of  $p$ . Or, under the terms of our account, the property X should be such as to guarantee (under the appropriate modal idiom) that the informant's beliefs whether  $p$  are correct, given the background conditions remain the same. Can we accept this?

Firstly, prior to considerations brought in by objectivisation, the inquirer cannot be interested in holding fixed all circumstances independent of R and  $p$ , but only the features of those circumstances which he himself is aware of. In the next chapter, we will see what happens to this reservation as objectivisation proceeds. Secondly, as with the above interpretation of Unger's non-accidentalism, I believe the absolute character of the formulation means that it is indeed too strong a condition to characterize what it means to be a good informant. To see why this is the case, we must return in the next chapter to the situation of the practical inquirer, to see just how demanding a typical inquirer will be. This is the subject of the next chapter, in which the argument against infallibilism about knowledge will be completed.

# 4. Interest-Relative Contextualism

## 4.1 Fallibilism and Practical Interests

Let us recap over the situation. Our practical inquirer finds himself situated in the world, and seeks to find out whether some proposition  $p$  is true – for the most part, in order to gain some practical advantage in the world. Moreover, there are a range of things he takes himself to know about the world, and a (much larger) range of facts about the world that he does not know. Consequently, his epistemic situation (understood in the narrow, internalist sense) will mark out a range of ‘open possible worlds’ that he could be in, given that his experience is a certain way. Any property  $X$  that is a serious candidate for correlating well with being reliable about  $p$  will show the informant to be right in their beliefs about  $p$  across a wide range of these possible worlds. Moreover, in a particular case, when the inquirer discerns that a potential informant is in possession of such a property, this will also determine an epistemic probability (from the point of view of the inquirer) that the potential informant’s beliefs about  $p$  are correct. Which of these two formulations is the most fundamental – and how demanding is our inquirer to be here?

According to the infallibilist, the answer to the later question is that the inquirer should be in some sense maximally demanding – so that, given the property, the truth of the subject’s beliefs is entailed (either logically or nomically) to be correct. This given, there is no difference between the modal and epistemic formulations – if we know someone is right in all the possible worlds that might be the actual world, then our epistemic probability of their being right is 1 (and conversely).<sup>4</sup> However, as we have seen, when we take the purpose of the concept of knowledge as determining its core features, and not a side-issue merely relating to pragmatics, we see that the infallibilist’s demands will not in general be met: for one thing, there are some possible worlds consistent with the inquirer’s experience that will be paid no attention (such as worlds where BIV and other radical sceptical scenarios obtain). Moreover, to express things probabilistically, I will argue below that it will only rarely be the case that when seeking an informant for some practical issue one will only settle for one that is likely to be correct with (epistemic) probability 1.

Given that the two notions come apart, then – to say something holds in a wide range of open possible worlds is by no means to say it is (epistemically) likely – I will suggest that either a modal or a probabilistic deficiency in an individual’s credentials is enough to undermine their prospective claims to knowledge whether  $p$  (as we shall see in section 4.5 below, this echoes a suggestion of Jason Stanley – Stanley, 2005). In the previous chapter, we saw how the third person nature of our enquiry and the consequent set-up of the prototypical case meant that for some kinds of questions (i.e., practical questions about the imminent world) some open possible worlds (e.g., BIV worlds) were likely to

---

<sup>4</sup> Mathematicians distinguish between ‘certain’ and ‘with probability 1’, with the former entailing the latter but not conversely. However, for our purposes this distinction may be safely ignored.

always be excluded from consideration in determining when our anthropomorphized concept of knowledge was applicable. Later in this chapter I will explore precisely which sets of possible worlds an inquirer may exclude; in the remainder of this section I will explore how epistemic probabilities less than 1 might come to be acceptable.

If I am enquiring for my own purposes, I will attach a certain level of importance to gaining correct beliefs whether  $p$ . The possibility that an informant might be wrong is weighed against various other facts, such that any belief about  $p$  might be better than no belief: the real world has its own time pressures, and sometimes even an action based on incorrect evidence is better than no action at all. Moreover, there may be some negative consequence to not coming to know whether  $p$ . For instance, suppose I wish to know whether a given tree has bananas at the top of it; I cannot tell from the ground, but I am aware that Peter climbed the tree the other day (a property that makes him reliable on the question at hand, but not infallible; someone may have interfered with the tree since). My practical interest in acquiring bananas to eat, the amount of effort that I will need to expend climbing the tree, the opportunity costs waived in embarking on this adventure, and my attitude to risk all contribute to determining some minimal level of epistemic probability that must be met in order for it to be rational (or strategically prudent) for me to follow the information given to me by Peter. Unless I am maximally risk-averse – and there is no reason to suppose individuals in the state of nature would be like this; in fact, everything we have said so far seems to speak against it – then this probability will usually be less than 1.

In the general case, whenever I am enquiring for my own purposes, I will take all of the practical facts of this nature (of which I am aware) into account and adopt a threshold of ‘likeliness to be correct’ that must be met in order for an informant to count as acceptable to me – that is to say, I will only ask informants who appear to possess some property  $X$  which suggests that they are (epistemically) at least *this* likely to be correct on the matter in question. This is Craig’s condition I4 that we met in chapter 2. However, as was also made clear in chapter 2, we have thus far considered only a simplified version of things: knowledge ascriptions are not made in isolation, but in general in our linguistic community they will be made publicly, and hence have certain repercussions. Likewise, sometimes I may want to recommend an informant to Fred, without having any idea what Fred’s purposes are in wanting to know – so, one might think, I had better make sure that I recommend someone that is sufficiently likely to be correct *whatever* Fred’s purposes might be (i.e. correct with probability 1). However, if this is our strategy we are likely to find ourselves usually depriving others of access to informants that are in fact perfectly sufficient to their needs.

Later in this chapter (section 4.4) we will explore more fully the effects that these kinds of contrary pressures objectivisation leads to will have on the threshold of probability necessary for knowledge; meanwhile, in the next two sections we return to the modal aspect of the debate. In the next section we will look to the literature for an attempt to capture the sets of open possible worlds that are salient to knowledge ascriptions: the contextualism of David Lewis (Lewis, 1996).

## 4.2 Lewisian Contextualism

We saw in the last chapter that the infallibilist's formulations of the concept of knowledge tend to involve modal conditions that are too strong, and hence looser conditions are necessary. But how is this 'looseness' to be expressed, exactly? Roughly speaking, what is required is to demarcate some non-exhaustive subset of possible worlds such that it is necessary for knowledge that the detectable property X ensures an informant is correct across all of them (or, alternatively, some larger subset such that the informant is correct across most of them – including the actual world). Let us try out a proposal of this sort by looking at a recent attempt to introduce some flexibility into the concept of knowledge: the contextualism of David Lewis. An utterance is context-sensitive if it expresses different propositions in different contexts of use; contextualism is the thesis that knowledge ascriptions are context-dependent in 'a distinctly epistemological way' (Stanley, 2005). This means that in different contexts of use, different evidential standards are called into play. Lewis's proposal is that:

**S knows that P iff S's evidence eliminates every possibility in which *not-P* - Psst! - except for those possibilities that we are properly ignoring. (Lewis, 1996, p554)**

Again, adapting the proposal to our overall account of knowledge, we may understand this as a proposed necessary condition on the property X – that possession of it excludes every possibility in which the subject is incorrect about *p*, except those we are 'properly ignoring'. Moreover, Lewis supplies a list for conditions for when we are and are not properly ignoring a given possibility. Briefly, these are:

### **The rule of Actuality**

The possibility that actually obtains (i.e. the possible world we are in) is never properly ignored.

### **The rule of Belief**

A possibility that S believes to obtain is never properly ignored.

### **The rule of Resemblance**

Supposing two possibilities saliently resembling each other; then if one may not be properly ignored (other than in virtue of this rule), then neither may the other. For example, the possibility that you win the lottery saliently resembles the actualized possibility in which you do not, and so cannot be properly ignored.

### **The rule of Reliability**

Where we obtain knowledge by a route that is in general reliable (such as visually), the possibility that this mechanism has suddenly failed us on this occasion may properly be ignored.

### **The Two rules of Method**

We are entitled to presuppose that a sample is representative, and that the 'best' explanation of our evidence is correct.

### **The rule of Conservatism**

Possibilities which are usually ignored by others around us may be ignored.

### **The rule of Attention**

A possibility that is not actually ignored [i.e. to which our attention is directed] is not properly ignored.

Lewis later asserts that there is no reason to expect there to be a purpose that the concept of knowledge thus demarcated is suited to fulfilling: ‘What is it all for? Why have a notion of knowledge that works in the way I described? (Not a compulsory question. Enough to observe that we do have it.)’ (Lewis, 1996, p 563). Of course, given the kind of project I have embarked upon, I should have to respectfully disagree here. He goes on to say: ‘if you doubt that the word ‘know’ bears any real load in science or in metaphysics, I partly agree.’ Indeed, if knowledge is characterized in this way, it becomes hard to see any role for it, primarily because the rule of attention. This rule implies that, should an individual consider whether they have hands (for example) whilst bearing in mind the possibility of being deceived by a Cartesian demon, then they do not know whether they have hands. But if they are not considering such a possibility, then they do know (Lewis wants to say). However, as long as a potential informant meets the other conditions laid down for knowledge in this thesis, it should be of no concern to an inquirer whether the subject is consciously turning over this possibility or not; the property of being a good informant is more substantial than that. The fact that they are merely considering such a possibility does nothing to undermine their claims to knowledge (so long as they do not take it so seriously that they become distracted and thus unable to properly tell us anything).

On the other hand, at least some of the other rules may plausibly be thought to have motivations from the state of nature perspective. Consider, for example, the rule of conservatism; if everyone ignores a possibility it is unlikely to feature in our constructed concept as actually operated with when analysed from the descriptive or strategic frameworks. Moreover, given also that such possibilities are actually false, it might be rational to ignore them too, especially if not doing so excludes large classes of informants on this basis alone (potential informants may not bother ensuring that their beliefs *would* be correct in worlds they are sure are not the actual world). Likewise, the ‘rule of actuality’ reiterates the factive condition of knowledge, and the ‘rule of belief’ will usually be met wherever we can properly regard a subject S as (truthfully) ‘telling us whether *p*’, given that people in general try to avoid holding contradictory beliefs. The rule of reliability has an analogue with the discussion from chapter 3 whereby a reliable connection between X and having true beliefs whether *p* is (given the other conditions are met) sufficient for knowledge, even though there are conceivable circumstances where this connection could fail. And as the rule of resemblance seems intuitively plausible, perhaps we can find a place for this too (the inquirer may not know which of these alternatives bearing close resemblance to each other are actualized).

In general, if our hypothesis is correct, then if there are respects in which knowledge is context-sensitive, they must have something to do with the context-relativeness of being a suitable person to use as an informant as to whether *p*. These in turn will tend to relate to either the *general* circumstances surrounding enquiry in our primitive society, or to the practical interests of the inquirer. In seeing where Lewis’s conditions diverge from these criteria – chiefly with the rule of attention – we can see where his analysis appears – from the Craigean perspective – to have gone wrong; moreover, considering where he seems to have got it right will help steer us in the right direction.

### 4.3 Possible Worlds and Objectivisation

Let us recap over the simplest case of knowledge ascription. In the prototypical scenario, an inquirer will flag an external individual as a good informant, based on some detectable property *X* that suggests they will be correct in asserting whether *p* (where *p* is some practical, everyday matter) across some range of possible worlds that the inquirer considers, for these purposes, to be open. The sense of ‘consider’ here means to take into account in his or her practical actions; this is not Lewis’s ‘rule of attention’: there are, more or less, a fixed set of possibilities we ignore when applying the concept of knowledge in everyday situations, and these generally include the favourite possibilities of the sceptic. Some of these possible worlds, such as where the inquirer is a BIV, are excluded by the terms of the enquiry. Moreover, the third-person nature of the ascription means the inquirer will not consider some worlds that are consistent with how things seem to the potential informant – such as worlds where the potential informant is a BIV. If this individual is correct across the possible worlds considered, including the actual world, we may say that they are a suitable informant for this inquirer to use, and hence appropriate subjects for knowledge ascription according to our hypothesis.

Now, for conditions I1 – I3, the effect of objectivisation of our constructed concept was to take us from relativised conditions that depended on the situation of the inquirer, to more substantial, intersubjective conditions. The thought here is that perhaps the same process will take us from this range of open possible worlds that depend on the inquirer’s situation, to a more general condition simply referring to possible worlds centred around the informant that are close to the actual world – hence ending up at an account with modal aspects somewhat similar to that of Robert Nozick (Nozick, 1981).

Consider two inquirers in different epistemic situations, each considering whether to use an individual as an informant. Each brings something different to the table; the range of possible worlds they consider to be open will be different (but overlapping), and they will have varying degrees of information about the informant herself (indeed, it might be a completely different property *X* that each is using to identify them). We may leave the second kind of consideration aside, because this is information that inquirers can easily and swiftly compare (i.e., we may consider what would happen if everyone had first shared all of the information of this nature). With condition I2, the detectability requirement, the effect of objectivisation was to approve knowledge ascriptions across the *union* of the set of informants recognizable to each observer; we thus abstracted away from the detection capacities of particular individuals to a non-relativized, general notion of detectability. What is the equivalent process of combination for condition I4?

Firstly, if everyone in the society does not take seriously the possibility that some particular class of possible worlds obtains, then this class of possible worlds will not be considered when evaluating good informants (c.f. Lewis’s ‘Rule of Conservatism’). In other words, knowledge ascriptions occur against a fixed backdrop of ‘common knowledge’ within the society. Hence, it will be permissible for knowledge-ascriptions relating to informants that are correct across the union of open possible worlds of every competent observer in the society (or in some restricted sense, every *possible* observer) to enter into public circulation. Should we stop here?

Supposing I were to recommend an individual as a good informant on some particular question, because I can tell he is correct in all of the possible worlds I seriously consider

to be open. Should it concern me that the person to whom I am recommending considers other possible worlds to be open; worlds where the informant might not be correct? Again, we have two contrary pressures; on the one hand, no; I can in good conscience recommend this person, and allow this recommendation to become commonly believed, because I firmly believe that the informant's beliefs on the matter in hand will be correct. It is true that if another inquirer seriously considers some possible world to be a possibility, then if the potential informant will not be correct in that world, the inquirer will not think he is a good informant. But from my point of view, this might be diagnosed as a case where, due to his limited perspective on the situation and what possible worlds might be actualized, this other inquirer is simply wrong. On the other hand, however, I would also like these recommendations to be believed: firstly due to the altruistic motive of wanting to enable others to make use of them, and secondly because presumably it is in my interests not to acquire a reputation for being inept at recommending informants (this second desire might be partly motivated by the first), or to suffer reproach for recommending someone inappropriate.

It seems the situation here is like this: we can imagine two inquirers, each having a different set of open possible worlds, and we are interested in which set is the correct one to consider for knowledge ascriptions. Now, it is likely that there is no objective answer here: only one world is *really* open, given all available information, and that is the actual world. But if we restrict consideration to this world alone, we should be left with no modal condition at all – and we will have clearly missed the point somewhere.

Perhaps it would be better to describe the situation like this: there are general background assumptions that people in our primitive society share, and only possible worlds where these assumptions hold true are relevant to knowledge ascriptions (these worlds are thus in some sense 'close', although not expressly understood by inquirers as centred around the actual world, for they do not know which of these open possible worlds this actual world is). However, these assumptions are neither static nor overly determinate and explicit: given that we might be in a privileged epistemic position with regard to what possible worlds might actually be the real world, we can also make recommendations to others. If a recommendation is sufficient to persuade the person to whom we are recommending that the open (from their perspective) possible worlds in which the subject are wrong are in fact not candidates for actuality, the recommendation will be successful, and it will be in our interests to make it. If not, we always have the option of simply asserting that we believe their beliefs on the matter in question will be true – possibly supplemented with some kind of explanation as to why. Moreover, if the extra information we supply in excluding some of the worlds in which the potential informant would be wrong is overly specialized, detailed or technical, we then cease to be able to regard this potential informant as someone who is properly able to 'tell us whether *p*'.

#### 4.4 Objectivisation and Probability

Let us now return to the probabilistic aspect of the debate. We have said (pre-objectivisation) that our inquirer seeks an informant who is 'likely enough to be correct for his [the inquirer's] purposes': if the question is one such that even an answer with a good chance of being wrong is better than no answer at all – perhaps we are subject to

some penalty if we do not act, such as starvation – then we may be rather more lax in our standards for considering someone a good informant. However, if our purposes in wanting to know are very important indeed, we will tend to be much more discriminating when we come to decide whom we are willing to trust. This given, in the simple, individual case it seems that the probability threshold required – the minimum epistemic probability with which the detectable property  $X$  must predict a correct answer about  $p$  – will be relativised to the inquirer's concerns. In the next section we will consider (Stanley's) objections to the claim that knowledge is context-dependent in this way; in this section we will consider what happens to this condition as objectivisation proceeds.

As previously noted, I may often find myself in a position to recommend, perhaps via a mediator, informants to an inquirer, whilst having no information about what that inquirer's purpose is in wanting to know (other than what is suggested by the nature of the question). This given, there is something of a 'pull towards certainty' (in Craig's words) here – perhaps I should do well to only recommend informants who are likely enough to be correct to meet any practical purpose this informant might conceivably have. The same is true if knowledge ascriptions are made with no particular purpose in mind, but simply to add to the general cumulative store of information that is shared amongst the community: we do not know the purpose for which they will be used.

We may supplement the thought that there is something of a pull towards demanding increasingly competent informants in our ascriptions of knowledge with a notion frequently discussed in the literature: that of a lottery. Whilst certainly not an uncontroversial feature of the human condition, it is plausible that lotteries and other forms of betting at quantified rates are widespread enough to bear influence on current everyday conceptual practice in the world of today (we can thus consider this an extension of the state of nature method). What the lottery example provides is an everyday situation in which an inquirer's needs may be made arbitrarily demanding: if a lottery has  $n$  tickets, then the testimony of an informant who claims – with probability less than  $n-1/n$  of being right – that the inquirer will lose if he buys a ticket may find that his testimony is ignored if the lottery's jackpot is sufficiently high.

Given that there are common situations where an inquirer may become arbitrarily demanding, then, as well as that oftentimes we will not know in advance what purpose a knowledge ascription may be used for, we might think that objectivisation will lead to a situation where the concept of knowledge is only correctly applied to informants who are certain to be correct. However, this line of thought is too quick: it assumes that the best conceptual strategy would aim at never being wrong. By setting our threshold at certainty we would achieve this goal, but in insisting upon this condition we also exclude from proper usage vast numbers of cases where a particular informant is perfectly suitable for the purposes an inquirer might have (bearing in mind specifically the case where this is so but not known to us). Our withholdings of knowledge have social consequences: in doing so we are suggesting to others that they do not act on this information; that they do not treat certain individuals as good informants. And when there is clear benefit in them treating them as good informants – the informant was in fact sufficiently likely to be correct, given their purposes in wanting to know, as well as factively correct – it seems we will have spoken incorrectly, given our assertion of the primacy of the role of the concept. This would therefore be a poor strategy for a community to adopt.

Moreover, there is also the possibility of error on our part – given the unpredictability of real-world conditions and the imperfections inherent in (for example) our sensory apparatuses, it does not make strategic sense to be this demanding. A potential informant will rarely, if ever, possess a property that makes him likely to be correct with probability 1 – and even if this is the case it will be nigh on impossible for us to identify that they do (as opposed to falling just short of the mark, say). This given, considerations such as the theoretical possibility of arbitrarily demanding inquirers will most likely bear little influence on conceptual practice. The pull towards certainty is offset by conflicting considerations (again, it is only because of the existence of contrary pressures that there is any controversy here at all).

What I suggest instead, then, is that there will be in operation a general level of probabilistic stringency for knowledge ascriptions (when one is blind to the purpose the recommendation will be used for) that will emerge from the balancing of the two sets of opposing forces – the pull towards certainty, and the pursuit of practical advantage. Moreover, rather than considering the epistemic probability for different inquirers, we can abstract to the conditional probability linking possession of the property  $X$  to being right about  $p$  (if there is more than one such property, as in general there will be, we consider the one corresponding to the highest conditional probability).

If a subject is suitable for a given inquirer but does not meet this general threshold level, it should not enter into common parlance that the subject ‘knows’ (we can always fall back on other expressions such as ‘it’s almost as if they knew’ – c.f. Craig, p98). However, even if an informant does meet the general stringency requirement, we should not rely on them if our specific requirements are more demanding, and likewise we should try not to recommend them to others if we are aware of their purposes in wanting to know and can see that this particular individual will fall short of their expectations. Similarly, given that this general level of stringency falls short of absolute certainty, someone with a particularly important and pressing reason for wanting to get the right belief on a given question is unlikely to rely on the general public opinion that a given individual does know the right answer (which doesn’t undermine the fact that this general opinion serves an important purpose).

Are these extra considerations relating to specific purposes merely pragmatics, or are they part of the core of the concept? We will take up a related linguistic question in section 4.6, below; for now we will continue to investigate proper conceptual practice, taking the role of the concept as primary. And as this condition has a clear justification from consideration of our hypothesis about this role, we may consider it as part of the core of the concept (where it applies). Hence, we should expect that knowledge ascriptions are often relativised to an inquirer’s concerns: to ascribe knowledge to someone, they must both meet the general level of stringency, as well as being likely enough to be correct for whoever’s purposes we have in mind when making the ascription (in those cases where we do have some such individual in mind at all).

#### 4.5 Stanley’s Interest-Relative Invariantism

The concept of knowledge as we have constructed it from our hypothesis bears some relation to another formulation found in the literature: that given by Jason Stanley in ‘Knowledge and Practical Interests’ (Stanley, 2006). They are similar in the sense that

both take the ‘practical facts’ – understood as ‘the costs of being right or wrong about one’s beliefs’ – into account (Stanley, 2006, p6). However, Stanley’s project is of a first person character, where knowledge is seen as a concept one applies in order to justify one’s actions being based on a particular belief (or, as we find in Williamson, as the ‘norm of assertion’ – Williamson, 2000). The emphasis is thus on practical facts relating to the subject of the knowledge ascription, whereas under our formulation it is the practical facts of whichever individual’s situation we are taking into account (of which the subject, and indeed the attributor themselves, are merely special cases). I should also point out that the analysis Stanley gives is meant only as a token attempt to encapsulate the concept of knowledge, the purpose of which is merely to demonstrate the manner in which practical facts are to be relevant to knowledge ascriptions. Stanley gives the following conditions for knowledge of  $p$  ( $x$ ,  $w$ ,  $t$  and  $p$  stand for names of person, worlds, times and propositions respectively):

**(1)  $p$  is true at  $w$**

**(2)  $\neg p$  is not a serious epistemic possibility for  $x$  at  $w$  and  $t$**

**(3) If  $p$  is a serious practical question for  $x$  at  $t$ , then  $\neg p$  has a sufficiently low epistemic probability, given  $x$ ’s total evidence**

**(4)  $x$  believes at  $t$  that  $p$  on the basis of non-inferential evidence, or believes that  $p$  on the basis of a competent inference from propositions that are known by  $x$  at  $t$ . (Stanley, 2006, p89/90)**

In defending this ‘interest relative invariantism’ against contextualism and relativism, Stanley refers to five cases, the last of which I shall quote in full:

*‘High Attributor-Low Subject Stakes.* Hannah and her wife Sarah are driving home on a Friday afternoon. They plan to stop at the bank on the way home to deposit their paychecks. Since they have an impending bill coming due, and very little in their account, it is very important that they deposit their paychecks by Saturday. Hannah calls up Bill on her cell phone, and asks Bill whether the bank will be open on Saturday. Bill replies by telling Hannah, ‘Well, I was there two weeks ago on a Saturday, and it was open.’ After reporting the discussion to Sarah, Hannah concludes that, since banks do occasionally change their hours, ‘Bill doesn’t really know that the bank will be open on Saturday’. (Stanley, 2006, p5).

with Stanley approving Hannah’s final utterance as seeming intuitively correct. There are four preceding cases that have similar structure and content: in ‘low stakes’, there is no impending bill, and Hannah attributes knowledge to herself about the bank’s opening hours, again based on having been there two weeks ago. In ‘high stakes’, Hannah admits that she doesn’t know whether the bank will be open, despite having the same evidence, because this time there is an impending bill. Stanley thinks that both these assertions are intuitively correct, but in ‘low attributor-high subject stakes’, an individual (Jill) – who has very little at stake in being correct – attributes knowledge to Hannah, and again it is important for Hannah to be right as there is a bill payment due. Stanley feels that this time the utterance is intuitively incorrect, even though the ‘low attributor’ could have ascribed knowledge to herself based on similar evidence. In ‘ignorant high stakes’,

Hannah and Sarah do have a bill due, but are unaware of it, and Hannah thus attributes to herself the status of knowledge – but Stanley claims that the intuitive reaction to this case is that she does so incorrectly.

Let us agree with Stanley's diagnosis of the appropriateness of knowledge ascriptions in each situation (given that the bank is in fact open on Saturday): how do these cases fit with the two (Stanley's, and that presented in this thesis) analyses? As Stanley argues convincingly in his book, his interest-relative invariantism gives the correct answer in each of the first four cases, but an incorrect answer in the final case (which we quoted first) 'High Attributor – Low Subject Stakes'. Stanley feels Hannah's final utterance in this case to be intuitively correct; however, Bill has nothing at stake (assuming his interests to be unconnected to Hannah and Sarah's) and hence Stanley's position suggests he *does* know, and that Hannah speaks wrongly. Stanley explains this discrepancy by appeal to a psychological pressure for us to project Hannah and Sarah's practical situation onto Bill; that is to say, not to ask whether Bill knows, but whether Hannah or Sarah would know if they were in Bill's situation.

On the other hand, the analysis I have presented seems to give the right answer in all five cases (so long as Hannah and Sarah meet the *general* stringency conditions in the first two cases – wherein if they do not then intuition has misled us here), because it is the practical facts relating to whomever the knowledge ascription is going to be made use of by that should be taken into account. If we allow ourselves to understand this 'projection' described by Stanley in a loose, metaphorical way – that consider relying on someone's evidence is similar to imagining what you would do in their epistemic situation – then this seems to coincide with Stanley's assertion about what is going on in 'High Attributor – Low Subject Stakes'. But then, we may simply disagree with Stanley about whether this projection is acceptable (and the intuitions appealed to seem to be on our side).

Later in the book, Stanley gives some independent arguments against contextualism – the thesis that knowledge ascriptions are context-sensitive (i.e. express different propositions in different contexts of use) in a distinctly epistemological way (i.e., the evidential standards properly in place depend upon the context of utterance). These are mainly linguistic in nature, and draw (primarily grammatical) disanalogies between 'know' and a wide variety of other context-dependent terms. We shall not have time to pursue these fully here, as our project is conceptual rather than linguistic. However, a common theme of these arguments is that if the term 'knowledge' were context-sensitive in the required way, then the evidential standards in play could change during a conversation in a way that made the utterances of the participants seem intuitively infelicitous. To take one of his examples:

- 'A. (Looking at a zebra in a normal zoo). I know that is a zebra.
- B. But can you rule out its being a cleverly painted mule?
- A. I guess I can't rule that out.
- B. So you admit that you don't know that's a zebra, and so you were wrong earlier?
- A. I didn't say I did. I wasn't considering the possibility that it could be a cleverly painted mule.' (Stanley, 2006, p52)

We have said that the evidential standards appropriate to knowledge ascriptions are relativised to the practical interests of the person whose interests we are taking into

account when making the ascription, if there is such an individual at all (oftentimes there will not be). But in a given situation, it is the (fixed) features of that situation that determine who we should properly consider that individual – the one whose interests we care about – to be: it is not a matter of where our attention is directed, but of a more rigid practical relation to potential action. Hence in our five examples it was Hannah and Sarah, because it is their actions that the knowledge ascription (either of another individual or themselves) will be used to justify or evaluate the appropriateness of (the veracity of a knowledge-ascription might therefore depend on who else was present at the time). The position I have advocated is therefore not contextualist in a manner that is damaged by these kinds of arguments, where knowledge standards shift within a discourse. If the account I have presented is correct, then it is to be expected that we cannot equivocate with regards to the individual whose interests are properly counted as most salient and still communicate successfully, at least without making explicit what is going on. Moreover, this equivocation would only occur if the participants' understanding of the situation and the interests and roles of the people in it were to change. Of course, we could construct a situation where this is the case – but if we are allowed to alter the participants' perceptions of the practical facts mid-dialogue then similar examples containing shifting evidential standards may be constructed for Stanley's position as well. Hence, our approach seems to have an advantage over Stanley's views, as it predicts the intuitively correct answers in all five test cases.

#### 4.6 Conclusion

We have thus completed our search for clarification of the conceptual practises surrounding the concept of knowledge, arriving at a kind of 'halfway house' between invariantism and contextualism. Whereas Craig (p166) seems to see fit to leave the probabilistic threshold connecting  $X$  and  $p$  as simply 'very high', perhaps specified by some qualitative condition such as 'justified in being certain', we allow for the standard to be raised when there is obvious need to do so – when consideration of the purpose for which the ascription will be used leads us to raising the standard to fulfil the needs of a particularly demanding inquirer. The conditions that encapsulate the prototypical case of conceptual ascription are therefore as follows. To be said to know whether  $p$ ,  $S$  must have:

**1. True belief as to whether  $p$**

**2. Various agent characteristics enabling him or her to properly 'tell us whether  $p$ '**

**3. A detectable property  $X$  reliably connected to being right whether  $p$ , such that:**

- i. The property entails the subject is correct in any possible world a typical attributor is seriously considering to be a possibility (given the general background assumptions in place in the society)**
- ii. The property entails the condition probability  $P(X|p)$  of the subject's being correct whether  $p$ , given that they possess  $X$ , is high enough to meet the general threshold in place in the community (i.e. for *most* everyday, practical purposes)**

**iii. If the knowledge ascription is made in a context where some particular purpose is involved (is connected to potential action), the conditional probability is suitably high for this purpose**

We have also seen knowledge ascription is intuitively withheld whenever we describe (from the global perspective) situations where the connection only holds by accident in this context, and the usual features making the connection reliable are absent.

In this thesis I have taken as a starting point the fact the concept of knowledge plays an important role in our lives, adopted Craig's hypothesis for what this role is, and used various frameworks to explore the features a concept suited to that role would have. These conditions therefore specify the cases where the concept is properly applied.

Moreover, the formulations of the infallibilists (having in mind here Unger and the exclusion principle) are such that the conditions I have put forth might be satisfied whilst their conditions are not. In resisting the application of the concept, we are resisting acknowledging that a potential informant is sufficiently qualified for the role: hence issues of practical advantage create contrary pressures that counterbalance the pull towards certainty, and towards the idealized conceptual practices that follow from adopting the infallibilist's absolutist concept of knowledge in our actual usage. Indeed, we have seen that it is this opposition of contrary forces that creates the tension in the concept of knowledge that enables such a fierce debate to exist at all (c.f. Craig, p 113). But, to return to the semantic question mentioned at the beginning of this thesis, do these conditions specifying when it is appropriate to attribute knowledge to the subject stem from the intrinsic semantics of the term 'knowledge' itself, or are they more like guidelines for appropriate deviation from its intrinsic semantics in actual usage (specifically with an eye to 3.iii)?

The contextualist suggests that knowledge ascriptions are literally correct whenever the purpose of the concept is met: when in doing so we are flagging a (partly contextually determined) suitable informant; the invariantist recognizes the practical advantage of making knowledge ascriptions where there is a chance the subject in question may be wrong, but claims they are, strictly speaking, not literally true.

In his book 'Philosophical Relativity' Peter Unger argues that the empirical facts of usage underdetermine the answer to this question; that given the complexity inherent in correct usage (for example, as characterised above), there is no principled way to apportion the complexity inherent in usage between the intrinsic semantics of the term 'knowledge' and pragmatic considerations pertaining only to how the term is (properly) used (Unger, 1984). But whilst this might be true, we have nevertheless achieved the purpose we set out with: to clarify the linguistic practices surrounding the concept of knowledge. If we were correct in claiming that knowledge as a phenomenon is delineated by the usage of its concept then it is these conditions of proper application that contain the interesting questions, and not the issues relating to how this is explained in terms of the underlying semantics. As Craig puts it:

'My 'practical explication' or 'state of nature' method leads to an account of the linguistic practice surrounding the word 'know' and its near relatives; it does not determine how we are to apportion the underlying mechanics of the practice between invariant semantics and contextually motivated pragmatics'.

If Unger's relativity thesis is true then this is the best we could have hoped for; if not, then the linguistic question is simply something we are not required to address (and, it seems, not something that this style of explanation is able to address).

Having thus reached the end of the project, we can also see that there is great scope for further research using the same methodology. One line of enquiry to be pursued would be a state of nature-style investigation into the features we should expect a concept suited to the (related) role of being the 'norm of assertion' should have – or the role of providing an underpinning or justification for one's actions: this is the first-person project mentioned in chapter 2. Indeed, there are not just one but a family of knowledge-like concepts, each bearing certain 'family resemblances' to each other in their application and the role they fulfil (c.f. Kusch, 2011, who provides a link between Craig's work and that of the later Wittgenstein). Craig discusses constructions relating to several more in his book – knowing Fred, which, further to simply knowing certain propositions about Fred, entails some kind of capacity to interact with Fred – and knowing London, and German (Craig, section XVI). Another project would be to strengthen the methodology by a more detailed empirical account of conceptual acquisition.

That there are further lines of enquiry open does not show that the project I set out to undertake has not been properly completed, and of course the same goes for Craig. Rather, this is what one should expect from this kind of naturalistic endeavour, as one would hope it to inherit aspects of research procedure from the natural sciences, where research efforts of one individual may add to those of another, rather than replacing them.

## 5. References

- Armstrong, D. 1973. *Belief, Truth and Knowledge*. London: Cambridge University Press.
- Craig, E. J. 1986/7. 'The practical explication of knowledge'. *Proceedings of the Aristotelian Society* 87:221-226
- Craig, E. J. 1990. *Knowledge and the State of Nature*. New York: Oxford University Press.
- Descartes, R. 1641. *Meditations on First Philosophy*. In Cottingham, Stoothoff and Murdoch, trans. 1984. *The Philosophical Writings of Descartes*. Cambridge University Press.
- Dennet, D. 1989. *The Intensional Stance*. 1989. Cambridge, MA: MIT Press.
- Dretske. 1971. 'Conclusive Reasons'. *Australian Journal of Philosophy*. 49: 1, 1-22.
- Goldman, A. 1967 'A Causal Theory of Knowing'. *Journal of Philosophy* 64: 357-372.
- Kusch, Martin. 2011. 'Knowledge and Certainties in the Epistemic State of Nature'. *Episteme* 8:6.
- Lewis, D. 1996. 'Elusive Knowledge'. *Australian Journal of Philosophy* 74: 549-567.
- Nozick, R. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
- McGinn, C. 1984. 'The concept of Knowledge'. *Midwest Studies in Philosophy* 9(1): 529-554.
- Pfeifer, R. and Bongard, J. 2007. *How the body shapes the way we think*. Cambridge, MA: MIT Press.
- Ryle, G. 1949. *The Concept of Mind*. Hutchinson.
- Radford, C. 1966-67. 'Knowledge—by examples'. *Analysis* 27: 1-11.
- Stanley, J. 2005. *Knowledge and Practical Interests*. New York: Oxford University Press.
- Unger, P. 1968. 'An analysis of factual knowledge'. *Journal of Philosophy* 65(6): 157-170.
- Unger, P. 1971. 'A Defense of Skepticism'. *Philosophical Review*: 198-219.
- Unger, P. 1975. *Ignorance—a Case for Scepticism*. Oxford: Clarendon Press.
- Unger, P. 1984. *Philosophical Relativity*. Oxford: Basil Blackwell.
- Valberg, J. 2007. *Dream, Death and the Self*. Princeton, New Jersey: Princeton University Press.
- Williams, B. 1978. *Descartes: the Project of Pure enquiry*. Penguin.
- Williams, B. 2002. *Truth and Truthfulness*. Princeton, New Jersey: Princeton University Press.
- Williamson, T. 1998. 'The Broadness of the Mental: Some Logical Considerations'. *Noûs* 32 (12): 389-410
- Williamson, T. 2000. *Knowledge and its Limits*. New York: Oxford University Press.