The Epistemic Justification Puzzle

Christos Kyriacou

A thesis submitted for the degree of Doctor of Philosophy to

Philosophy Department

School of Philosophy, Psychology and Language Sciences

The University of Edinburgh

2010

With love and gratitude,

to my parents Sophocles and Militsa

Declaration

I hereby declare that this thesis is of my composition and that it contains no material previously submitted for the award of any other degree. The work of this thesis has been produced by myself -except where due acknowledgement is made in the text.

Christos Kyriacou

Abstract

The thesis explores the semantics of epistemic justification discourse, a very important part of overall epistemic discourse. It embarks from a critical examination of referentialist theories to arrive at a certain nonreferential, expressivist approach to the semantics of epistemic justification discourse. That is, it criticizes the main referentialist theories and then goes on to argue for an expressivist approach on the basis of its theoretical capacity to outflank the problems referentialist theories meet. In the end, I also identify some problems for a prominent expressivist theory and, as a response to these problems, propose a novel norm-expressivist approach that seems to evade these problems.

In particular, in Ch.1 I introduce what I call 'the epistemic justification puzzle' and then in Chs.2-4 criticize naturalistic referential theories: analytic naturalistic reductionism, synthetic naturalistic reductionism and epistemic kinds realism. In Ch.5 I criticize nonnaturalist referential theories: what I call 'naïve' nonnaturalism and J.McDowell's (1994) more sophisticated quietist version of nonnaturalism. Next, in Ch.6 I introduce the semantic programme of expressivism and go on to construct a simple version of epistemic norm-expressivism (inspired by A.Gibbard (1990)) in order to explain how expressivism can easily outflank the identified problems of referentialist theories. This simple norm-expressivist theory, however, is only used as a theoretical 'toy' for the mere sake of motivating the possibility of expressivism, as in Ch.7 I go on to argue for a more sophisticated version of norm-expressivism: habits-endorsement expressivism.

In Ch.7 I introduce a prominent expressivist theory of moral and knowledge discourses, namely, plan-reliance expressivism (credited to A.Gibbard (2003, 2008)) and extend it cover the epistemic justification discourse. I then identify some problems for plan-reliance expressivism as extended to cover justification discourse and in response to these problems propose habits-endorsement expressivism. Habits-endorsement expressivism builds on the intuition that (justified) belief-fixation is habitual and exploits the theoretical flexibility of the notion of habits in order to overcome the identified problems of plan-reliance expressivism.

'Now, what do you think justification means? '

<div align="right">Plato, *Theaetetus* 206c7-8</div>

'I don't know yet; we must go wherever the wind of the argument carries us.'

<div align="right">Plato, *Republic* 3, 394d8-9</div>

Preface and Acknowledgements

This thesis is about the relatively nascent field of meta-epistemology, namely, the field inquiring into the second-order issues underlying first-order substantive epistemology. Such second-order issues involve questions in meaning, mind, (meta)epistemology, ontology, intuitions, logic, objectivity, (dis)agreement and even more. In particular, the thesis focuses on the epistemic justification discourse, an important part of overall epistemic discourse, and explores the semantics of epistemic justification assertions and attributions. That is, what we mean when we assert and attribute justified beliefs.

The overall dialectical plan is to expose the problems of prominent referentialist approaches and then explore a nonreferential, expressivist approach to epistemic discourse that promises to fare better. If, indeed, such a nonreferential approach can fare better, then we can make sense of the epistemic justification discourse without the postulation of any epistemic properties (and the concomitant problems they incur). Yet, the reader should be warned from this early point that as the literature on epistemic justification is vast, in order to make the fulfilment of my dialectical plan possible, I often had to privilege breadth of scope rather than depth of analysis. Unavoidably, not every single theory has been discussed and even the prominent ones that have been discussed, they have not been discussed to the substantial depth they surely deserve.

Table of Contents

Part I The Problems of Epistemic Referentialism

Chapter 1 The Epistemic Justification Puzzle

1.1 Unpacking the Puzzle

The aim of this work is to pursue an inquiry into the semantics of one of the key parts of epistemic discourse: epistemic justification discourse. It is to pursue an inquiry into what we mean when we think and say that a belief, a theory, an explanation etc. is (epistemically) justified. That is, what we mean when we think and say things like 'S justifiedly believes that p', 'S's theory is justified' and the like[1].

Unsurprisingly, as any other philosophically interesting notion, epistemic justification gives rise to a puzzle as soon as we realise the relevant background assumptions in play. That is, it gives rise to a number of individually plausible premises which are, unfortunately, jointly inconsistent. This is what we may call *'the epistemic justification puzzle'*[2]:

P1 The Property Assumption: All properties are either natural or nonnatural.

P2 The Referential Semantics Assumption: Justification assertions and attributions purport to refer to a certain property (natural or nonnatural) of epistemic justification.

P3 The Naturalism-Denial Assertion: Justification assertions and attributions do not purport to refer to a certain natural property of epistemic justification.

P4 The Nonnaturalism-Denial Assertion: Justification assertions and attributions do not purport to refer to a certain nonnatural property of epistemic justification.

The starting point of the puzzle seems to be the -prima facie innocent- Socratic question '*What is Epistemic Justification?*'. Once we pose this question, though, the seas of argument raise high and 'the epistemic justification puzzle' pops to the

---

[1] In what follows, I will mostly be concerned with what it takes for a belief to be epistemically justified, though, parallel considerations hold for other objects of epistemic justification like theories, explanations and inferences.

[2] This might not appear at first sight a 'puzzle' to some people either because they might accept all the premises of the puzzle (e.g. error theorists, relativists) or because they find some of the premises utterly implausible (e.g. nonnaturalists the P4). If that is the case then the reader should see this formulation as a merely crisp way to kick off the argument. Why I am not really worried with error theory and relativism that seem able to satisfy all premises of the puzzle is briefly explained in section 1.4.

surface. Let me explain. To pose such a Socratic 'What is Fness?' question, it is to ask for the essential property constitutive of the nature of the notion in question. It is to search for a reduction of epistemic justification in terms of necessary and sufficient conditions. Indeed, this is the theoretical approach of '*Naturalistic Reductionism'*. Naturalistic reductionism endeavours to reduce epistemic justification to other non-epistemic, natural properties.

There are two ways to reduce a higher-order, epistemic property to a lower-order, non-epistemic natural property: the traditional analytic way and the relatively recent synthetic. *Analytic naturalistic reductionism* searches for a reduction of epistemic justification by traditional a priori semantic analysis. In the same way we reduce by semantic analysis, let us say, 'spinster' to 'old unmarried woman', 'bachelor' to 'unmarried adult man' or 'vixen' to 'female fox' we should attempt to reduce by semantic analysis the notion of epistemic justification (and other normative notions); or at least this is what this philosophical programme urges.

In contrast with reduction by traditional a priori semantic analysis, recently arrived *synthetic naturalistic reductionism* searches for a reduction in the pattern of the Kripke-Putnam a posteriori reduction of natural kinds and other science-theoretic terms[3]. In the same way we take advantage of the empirical results of scientific inquiry in order to reduce natural kinds like 'water' to 'H20 molecules', and science-theoretic terms like 'genes' to 'DNA molecules' or 'temperature' to 'mean molecular kinetic energy', we should attempt to reduce epistemic justification (and other normative notions); or at least this is what this philosophical programme urges.

Unfortunately, however, both analytic and synthetic naturalistic reductionism do not seem to be especially promising. On the one hand, as we shall see in Chs.2-3, reducing the property of epistemic justification by a priori semantic analysis proves to be a Sisyphean task, a task that we endlessly strive to accomplish but we never seem to really succeed. Despite the existence of a plethora of proposed analyses, none seems to be in a position to evade counterexamples and establish itself as a successful analysis of epistemic justification. For a semantic reason we will make explicit in the next Ch.2, the notion of epistemic justification (and other normative notions) seem to stubbornly resist successful analysis[4].

---

[3] See S.Kripke (1981) and H.Putnam (1975, 2004).
[4] Compare W.Alston (2005:11): 'the perennial quest [for an analysis of epistemic justification]... is

On the other hand, as we shall see in Ch.4, reducing epistemic justification (and other normative notions) in the a posteriori pattern of Kripke-Putnam identities does not seem to fare any better. Normative properties do not seem to be reducible to such a posteriori identities. Indeed, some would add that we don't even know how to begin for such an a posteriori reduction since, unlike natural kinds and science-theoretic terms, normative properties do not seem to have anything to do with the empirical, natural realm explored by scientific inquiry[5].

Or at least they don't have anything apparent to do with the natural realm, as we don't find normative properties like justification or goodness in nature in the same way we find natural kinds like water and salt. Thus, the constant failure of naturalistic reductionism (analytic or synthetic) motivates the plausibility of the intuitive *'Naturalism-Denial Assertion'* (premise 3 of the puzzle): Justification assertions and attributions do not purport to refer to a certain natural property of epistemic justification.

Once naturalistic reductionism (analytic or synthetic) appears unpromising, the way opens for *'anti-reductionist approaches'* to epistemic justification. There are two such antireductionist approaches (broadly construed): (a) *'Nonnaturalism'* and (b) *'Expressivism'*. First, nonnaturalism suggests that the property of epistemic justification stubbornly resists reduction to a natural property due to the plain reason that there is no such natural property to be found. Granted the truism of 'the property assumption' (premise 1) stating that all properties are either natural or nonnatural, it appears then that epistemic justification is a nonnatural property.

According to nonnaturalism, when we make justification assertions and attributions, we purport to refer to the unanalysable, sui generis, nonnatural property of epistemic justification. Yet again, as we shall see in Ch.5, the nonnaturalist approach seems to meet serious ontological and epistemological problems that make its prospects of success appear bleak. Thus, the serious epistemological and ontological problems of nonnaturalism motivate the plausibility of the '*Nonnaturalism-Denial assertion*' (premise 4): Justification assertions and attributions do not purport to refer to a certain nonnatural property of epistemic justification.

---

quixotic of the same order as the search for the Fountain of Youth'.
[5] See, for example, J.Lenman (2003).

Second, as we shall see in Chs.6-7, we have expressivist anti-reductionist approaches. Expressivists - unlike naturalistic reductionist and nonnaturalist anti-reductionist approaches- are nonreferential approaches that reject '*the Referential Semantics Assumption*' (premise 2). As a matter of principle, expressivists are hostile to a 'referential semantic strategy' where the meaning of a sentence is supposed to be determined by what the subject-predicate of the sentence purport to refer to[6].

That is, if the subject refers to an individual (e.g. grass) that falls within the extension of the predicate (e.g. green), then the sentence satisfies its truth-condition and it is true (i.e. 'Grass is green' is true if and only if Grass is green). If the subject refers to an individual (e.g. my dog) that does not fall within the extension of the predicate (e.g. green), then the sentence does not satisfy its truth-condition and is false (e.g. 'My dog is not green' is true if and only if My dog is not green).

Instead, expressivists appeal to a 'use semantic strategy' where the meaning of a sentence is supposed to be determined by how the agents use the subject-predicate of the sentence in question. For example, 'Grass is green' means Grass is green, because this is how the subject-predicate of the sentence is used to mean, according to the semantic and syntactic linguistic conventions that govern English language.

In this use-theoretic semantic picture, normative reference is set aside, as some expressivists debunk reference and truth-conditions completely and appear to slide to some form of normative expressivist relativism, while others are not willing to give up precious reference and truth-conditions and attempt to derive them from the use of normative sentences[7]. Whether their attempt to rescue truth within such an antirealist-expressivist framework is successful remains contentious, as any theoretical approach to truth seems to be[8].

---

[6] This is the tradition of G.Frege (1997), early L.Wittgenstein (2001) and D.Davidson (1996), among others, and the one that is assumed in formal logic and semantics textbooks; see C.Howson (1997) and Kearns (2000). Formal logic and semantics just assume that semantics are referential (or 'descriptive'). For discussion of Frege's philosophy of language see D.Wiggins (1999) and C.Travis (2008).

[7] Expressivists like A.J.Ayer (1946), J.Austin (1962) and A.Gibbard (1990) thought that normative discourse is not truth-apt. But more contemporary expressivists like S.Blackburn (1998), M.Timmons (1999), A. Gibbard (2003) and R.Brandom (2000) attempt to rescue truth in such a use-theoretic framework. Blackburn, Timmons and Gibbard whom are ideational expressivists appeal to deflationism about truth in order to motivate what they call 'quasi-realism', namely, an antirealist framework that aspires to account for all the realist-seemings of normative discourse. Brandom is an inferentialist expressivist and follows a different, Sellarsian strategy. He attempts to rescue truth by appeal to the social game of 'giving and asking reasons' for our normative claims.

[8] For arguments to this effect see J.Knanvig (2007), T.Cuneo (2007) and M.Lynch (2008). Note that Kvanvig, Cuneo and Lynch argue only against ideational expressivists. For a critic of Brandom's

Expressivist approaches (broadly construed) split again into two categories: (a) Inferentialist expressivism and (b) Ideational (or Lockean) Expressivism[9]. As good use semanticists, inferentialists and ideational expressivists suggest that the whole project of searching for a property of epistemic justification (natural for naturalistic reductionism/nonnatural for nonnaturalism) is misguided because it rests on the mistaken assumption of referential semantics.

Contra the referentialists who take *reference* to be the master concept of the semantics of normative discourse (moral and epistemic), expressivists deny that reference can be the master concept for the semantics of normative discourse. Both the inferentialist and ideationalist expressivist agree that when we use normative sentences we don't purport to refer to a certain corresponding normative property. Instead, they contend that *use* should be the master concept of the semantics of normative discourse.

But despite the shared hostility to referential semantics and the shared adherence to use semantics, inferentialist and ideational expressivism have serious differences on core theoretical issues. A key difference is that they interpret differently the claim that use is the master concept for a theory of the semantics of normative discourse. On the one hand, inferentialist expressivists contend that the use of normative sentences indicates that *inference* should be the master concept of the semantics of normative discourse while, on the other hand, ideational expressivists contend that the use of normative sentences indicates that *expression* should be the master concept of the semantics of normative discourse.

Yet, in spite of their substantial differences, both inferentialist and ideationalist expressivists concur that commitment to 'the referential semantics assumption' is unwarranted. They argue that expressivist semantics can accept the intuitive premises 3 and 4 of the puzzle (the naturalism-denial and nonnaturalism-denial assertions) and avoid the respective problems referentialist approaches encounter (Moorean unanalyzability intuitions, epistemological and ontological problems, epistemic supervenience etc.).

---

inferentialism see C.Peacocke (2004).
[9] The origins of Ideational Expressivism can, arguably, be traced back to at least J.Locke's (1973) *An Essay Concerning Human Understanding* and D.Hume's (1986, 2005) *A Treatise of Human Nature* and *Enquiries.* Other notable Lockeans include J.Austin (1962), P.Grice (1989) and A.J.Ayer (1946). The origins of inferentialist expressivism are often traced back to the early American pragmatists like C.S. Pierce (1991), W.Sellars (1997) and Hegel.

The rejection of 'the referential semantics assumption' comes at a price, though. Expressivists reject 'the referential semantics assumption' (premise 2 of the puzzle) that motivates the projects of naturalistic reductionism and nonnaturalism and this seems to have some obnoxious semantic implications. Because of these obnoxious semantic implications, some philosophers think that the rejection of referential semantics for expressivist semantics comes to an unbearably high price[10].

The reason is that expressivism encounters problems (like 'the Frege-Geach' and 'truth' problems) in handling 'semantic facts' like validity, consistency, truth, objectivity and (dis)agreement that any theory of semantics must accommodate. These problems instigate the belief that expressivism is an unpromising approach to the semantics of normative discourse and, hence, we should instead remain committed to referential semantics that can easily accommodate such semantic facts.

So far, we have seen how the puzzle springs from the initially innocent Socratic question 'What is epistemic justification?' and what motivates the acceptance or the rejection of each one of the premises 2-4 of the puzzle. Last remains the least, perhaps, controversial premise: *'the property assumption'* (premise 1). The property assumption states what seems to be simply a truism, namely, that all properties are either natural or nonnatural. In other words, the nature of reality can be explained either by appeal only to natural properties or to both natural and nonnatural properties[11]. A natural property is an empirical property or, as is sometimes put after S.Shoemaker (2004) and D.Lewis (2004), a property with independent causal efficacy. As one might imagine, a nonnatural property is a non-empirical property or, alternatively expressed, a property with no independent causal efficacy[12]; some people question this assertion but I don't think they are right (see n.14)[13].

---

[10] For discussion of (ideational) expressivism's semantic problems see P.Geach (1972), J.Searle (1969), W.Sinnot-Armstrong (2007), M.Schroeder (2008). For attempted solutions see S.Blackburn (1984, 1993, 1998), A.Gibbard (1990, 2003) and M.Ridge (2006).

[11] This can be understood in terms of 'the problem of universals' and the perennial debate between Platonic realists and nominalists. The nominalist thinks that all existent properties are natural and the nature of reality can be adequately explained only by appeal to such natural properties (in various ways). The Platonic realist disagrees and claims that there must exist both natural and nonnatural properties because, in order to adequately explain the nature of reality, we need to postulate the existence of nonnatural properties. For the problem of universal see D.M.Armstrong (1980) and J.P.Moreland (2001).

[12] For the natural/nonnatural distinction and the various ways being glossed in literature see M.Ridge (2003). For naturalism see M.Timmons (1999) and D.Papineau (2009).

[13] Some philosophers like N.Sturgeon (2007) and R.Wedgwood (2007) have argued that normative properties have causal powers and therefore normative explanations are causal. Sturgeon, in particular, has engaged in dialectical sparring with G.Harman (2007) in a series of papers on this controversy. I

Now, let us retrace our steps. Premise 1 seems uncontroversial and premises 2-4 are individually intuitive but jointly inconsistent and this gives rise to the puzzle. In order to solve the puzzle, one of the premises 2-4 must be dropped and then surmount the serious difficulties that respectively ensnare each one of the *three routes* out of the puzzle[14]. These three routes out of the puzzle are the following:

- Drop 'the referential semantics assumption' (premise 2), opt for the nonreferentialism of either (a) inferentialist expressivism or (b) ideational expressivism and fight through their respective semantic problems.

- Drop 'the naturalism-denial assertion' (premise 3), opt for nonnaturalism and fight through its epistemological and ontological problems.

- Drop 'the nonnaturalism-denial premise' (premise 4), opt for naturalistic reductionism and fight through its constant failure to accomplish a reduction (analytic or synthetic) of the notion of epistemic justification.

With the puzzle introduced and the possible routes out of it now in sight, a good question at this juncture is which theoretical project is the most promising in overcoming its respective problems and solving the puzzle. If we could decide which one is the most promising, then obviously this is the solution we should pursue. Unfortunately, this is a very difficult question to be answered without first pursuing the projects themselves and seeing their theoretical virtues and resources in practice. As one might say, we want to swim (in the seas of argument) without having our legs wet.

Still, we need to venture an answer if we are to explore a possible route out of the puzzle. Granted that the circumstances demand a choice, my choice is to opt for the expressivist route 1 out of the puzzle. That is, drop 'the referential semantics assumption' (premise 2) and explore a nonreferential solution. In particular, I will go for an ideational expressivist solution (1 (b)) to the puzzle and not an inferentialist (1(a)) one because I am more sympathetic to ideational expressivism rather than inferentialist expressivism. Discussion of inferentialist expressivism won't show up in the limited space of this thesis, and will have to wait for another context. I will

---

can't touch this deep issue here but I don't see how normative properties can have independent causal powers if they are not natural properties. For the same controversy see C.McGinn (1999), A.Gibbard (1990, 2003), S.Blackburn (1993, 1998) and R.Shafer-Landau (2003).

[14] There are other routes out of the puzzle that I choose to ignore here: Error (or Fictionalist) theories, relativist theories etc. I explain the reasons for ignoring these two theoretical possibilities in Section 1.4.

simply assume that ideational expressivism is the most promising expressivist variant and henceforth with expressivism mean ideational expressivism[15]. I will reserve inferentialism for inferential expressivism.

Also, in exploring expressivism, I will also assume that expressivism can somehow tackle its serious problems with semantic facts. My intention here is not to resolve expressivism's semantic problems and vindicate expressivism as such, but to explore what could motivate an expressivist approach to epistemic justification discourse and how such an expressivist approach could best go. Thus, discussion of these semantics problems won't substantially appear in the thesis.

Of course, if our assumption proves wrong and expressivism cannot tackle its semantic problems, then expressivism as an approach to normative semantics should put to rest. But even if this unfortunate for this thesis result obtains, then our exploration won't be entirely worthless as it could stand as an account of the pragmatics of epistemic justification discourse. That is, of the mental content expressed by speakers in various contexts where they make justification assertions and attributions[16].

As I said, I find the rejection of standard referential semantics and the adoption of expressivist semantics the most *attractive* and *challenging* way for exploring a solution to the puzzle. Hence, this is the solution to the puzzle I will explore in the thesis. If some justification is needed why I find the expressivist solution to be intuitively the most attractive and challenging, the answer is to point to other intuitions I have[17]. Very briefly (and rather unfairly to the complexities these intuitions touch), here are three intuitions that incline me to think that expressivism is the most attractive possible solution. First, I intuitively sympathize with a methodologically naturalistic framework and expressivism comports nicely with such an ontologically parsimonious framework.

---

[15] For the case in favour of inferentialist expressivism over ideational expressivism, see M.Chrisman (2010) whom argues contrary to my assumed theoretical superiority of ideational expressivism over inferentialist expressivism.

[16] No doubt, the distinction between semantics and pragmatics is not an easy to draw, but for present purposes I conveniently sidestep this issue. For discussion of the issue see Z.Gendler Szabo (2008).

[17] I discuss a bit this methodological issue in my 'Naturalism and Normativity' (2008). Also, in the thesis I talk a lot of intuitions. With an 'intuition' I mean what many philosophers, after G.Bealer (1998: 201-239), call 'an intellectual seeming'. It should be made clear, though, that appeal to intuitions does not commit one to some form of intuitionistic epistemology. For a nice disentanglement of these issues see B.Williams (1993:93-101).

Second, I find the expressivist project to be a very fertile, flexible and resourceful project with some alluring theoretical virtues (explains Moorean unanalyzability intuitions, explains epistemic supervenience etc.). Third, I find at least some of the discontents of the other approaches to be difficult to be surmounted ('ontological queerness' and 'epistemological spookiness' of nonnaturalism, 'open question' arguments etc.). Finally, a justificatory intuition for being inclined to think of expressivism as the most challenging possible solution. Simply, it is because it questions the predominant 'referentialist paradigm' (in the Kuhnian sense) and swimming against the mainstream current (of tradition and arguments) is always challenging, albeit somewhat risky[18].

Of course, these justificatory intuitions cannot be pressed too far because they are themselves highly controversial and, therefore, need to be backed up by arguments. They need to stand before the tribunal of reason and in the light of arguments be adjudicated whether they are justified or not. Quite possibly, the controversies these intuitions touch run so deep that they cannot be even directly judged at first sight, but need to be judged at the end of the journey from the theoretical fruit they ultimately produce. That is, they need to stand as working hypotheses and be judged in view of their theoretical implications, adequacy of explanatory power, theoretical virtues like simplicity and elegance etc. Besides, these intuitions are just preliminary 'starting point' intuitions and we know very well that prima facie reasonable intuitions often turn to have unpalatable implications or even be straightforwardly mistaken.

But at any rate, even when our intuitions are proven deceptive they are proven to be so in virtue of other intuitions. One way or another, as T.Williamson (2004) has put it, 'intuitions are the boat we find ourselves in'. In other words, intuitions are not something that can be forced upon our psychology and, as a natural consequence, you either find yourself with certain intuitions or you don't[19]. Since these starting point intuitions are the boat I found myself in, unavoidably, with these I am bound to set sail through the roaring seas of argument 'the epistemic justification puzzle' raises. I will set sail for the philosophical exploration with these intuitions as my boat and -as Plato's quote at the beginning of the thesis says- let to be carried away wherever the wind of the argument leads.

---

[18] Here I emulate Brandom (2000) who often talks in terms of 'the referentialist paradigm'.
[19] As Max Weber (1982) has said 'intuition cannot be forced'.

The epistemic justification puzzle now introduced, this is what follows in the rest of this introductory chapter. In the next section 1.2, I outline the argumentative plan of the thesis. I explain in general terms what each chapter contains and also say a few things about the habits-endorsement expressivist approach I sketch in the end. Then in section 1.3, I briefly adduce some more motivation for exploring a form of expressivism for epistemic discourse. The motivation consists of the fact that the debates in ethics and epistemology seem to move in parallel. This 'ethics-epistemology parallel', among other considerations, fuels the intuition that expressivism could also be fruitfully applied from moral- where it has predominantly been applied- to epistemic discourse.

Afterwards, in section 1.4, I devote some time for short but important clarifications. I briefly clarify my stance in relation to some important issues that 'the epistemic justification puzzle' directly or indirectly touches. These issues, inter alia, concern the epistemic internalism/externalism distinction, the pragmatic/epistemic justification distinction, the doxastic voluntarism/nonvoluntarism distinction and more. Finally, in section 1.5, I summarize the argument of the chapter and prepare the ground for the imminent Ch.2.

1.2 The Argumentative Plan of the Thesis

In the first section, I introduced the epistemic justification puzzle. I explained how the puzzle arises from the seemingly innocent Socratic question 'What is epistemic justification?' and explained what motivates the acceptance or the rejection of each one of the intuitively plausible but jointly inconsistent premises 2-4 of the puzzle. Finally, I stated that my approach to the puzzle will be an expressivist one and very briefly cited some of the preliminary 'starting point' intuitions that incline me to explore the expressivist route out of the puzzle. In this second section, I outline the argumentative plan of the thesis, namely, explain what the content of each chapter will roughly be and also give some clue about the contours of the habits–endorsement expressivist approach to the puzzle I intend to tentatively explore. Let us turn to that.

In terms of general structure, the thesis can be divided into two parts. The first and larger part under the label of '*The Problems of Epistemic Referentialism*' is mostly critical in character (Chs.2-5). It criticizes referentialist solutions to the puzzle and

prepares the ground for the advent of my favourite habits-endorsement expressivist proposal. This mostly critical part is supposed to make explicit the chronic problems of the rival referentialist approaches and prepare the ground for the introduction of the expressivist proposal. The second part under the label '*The Possibility of Epistemic Expressivism*' is constructive (Chs.6-7).

The first constructive part comes in Ch.6, where I use a simple version of norm-expressivism as a 'toy' expressivist theory in order to exhibit how expressivism can avoid the referentialist's problems and thereby motivate an expressivist approach to the puzzle. But this simple norm-expressivist theory is only a theoretical 'toy' used to motivate expressivist semantics, as in Ch.7 I develop a more sophisticated version of norm-expressivism.

In Ch.7 I sketch the basic contours of an original version of norm-expressivism, namely, 'habits-endorsement expressivism' that exploits the versatile notion of habits. Habits-endorsement expressivism is motivated by means of arguing that Gibbard's (2003, 2008) later expressivist theory for knowledge discourse, namely, plan-reliance expressivism, if extended to cover the epistemic justification discourse, runs into difficulties that habits-endorsement expressivism can adequately address.

In more detail, in Chs. 2-4 I criticize naturalistic reductionism (analytic and synthetic) in order to motivate the plausibility of 'the naturalism-denial assertion' (premise 3). In Chs.2-3, I criticize analytic naturalistic reductionism. In Ch.2 I make explicit the theoretical commitments of analytic naturalistic reductionism, clarify what conception of epistemic justification we are after, present the conditions that an adequate reduction of epistemic justification by semantic analysis should satisfy, gloss the vexing natural/nonnatural distinction and, finally, introduce what I call 'the Moorean/Humean lesson'. That is, a dialectical lesson drawn from the epistemic versions of G.E.Moore's 'open question argument' and D.Hume's 'is/ought argument' meant to undermine the plausibility of analytic naturalistic reductionism.

In Ch.3 I step forward to apply 'the Moorean/Humean lesson' against prominent attempts to reduce by semantic analysis the property of epistemic justification. These reductionist approaches (coherentism, foundationalism, reliabilism etc.) are seriously undermined by this application. In Ch.4, I first explain how the so-called 'epistemic twin earth argument' stands as a modernized 'Moorean/Humean lesson', this time

undermining sophisticated synthetic naturalistic reductionism. Second, I discuss and reject G.Sayre-McCord's (1997a) interesting proposal of taking normative properties as 'normative kinds', that is, as properties irreducible to natural kinds, though, what the exact ontological status (natural or nonnatural) of such normative kinds is supposed to remain open.

In Ch.5, I criticize nonnaturalism in order to motivate this time the endorsement of 'the nonnaturalism-denial assertion' (premise 4) of the puzzle. First, against what I call 'naïve' nonnaturalism, I exploit its ontological and epistemological difficulties: queerness, epistemic access and epistemic supervenience. I argue that it meets the 'queerness' and 'epistemic access' objections and, in addition, that it seems to fail to satisfy the supervenience desideratum on epistemic discourse.

As I explain, any account of epistemic justification should satisfy 'the epistemic supervenience desideratum' and evidently 'naïve' nonnaturalism does not seem able to do so. What exactly is the epistemic supervenience desideratum and why is a desideratum for any theory of epistemic justification is explained there. Afterwards, I discuss J.McDowell's (1994, 1998) more sophisticated 'naturalism of second nature' that with a quietist twist attempts to disguise nonnaturalism in naturalistic cloth and, therefore, evade its ontological and epistemological difficulties[20].

The upshot of the critique of referentialist approaches (naturalistic and nonnaturalistic) in Chs.2-5 is the motivation of the plausibility of 'the naturalism-denial assertion' and 'the nonnaturalism-denial assertion' and therefore -by elimination- the rejection of the remaining 'referential semantics assumption' (premise 2) as a possible route out of the inconsistency of the puzzle. This task is undertaken, first, in Ch.6 where I motivate the possibility of an expressivist approach by means of showing how expressivism can evade the referentialist's problems. That is, explain Moorean 'open feel' intuitions, explain epistemic supervenience and avoid any extra ontological and meta-epistemological burden. To that effect, I use as a 'toy' expressivist theory a simple version of norm-expressivism, albeit to be substituted by a more sophisticated one in Ch.7.

Second, in Ch.7 I constructively introduce a more sophisticated version of norm-expressivist approach to 'the epistemic justification puzzle', namely, habits-

---

[20] Of course, this is a dialectical intention I attribute to McDowell and not one that he explicitly asserts or one that he would be glad, I think, to embrace,

endorsement expressivism and sketch some of its basic features. I motivate habits-endorsement expressivism by means of arguing that plan-reliance expressivism (as applied to epistemic justification discourse) faces problems, that is, it runs into direct counterexamples and also defies 'the nonvoluntarism intuition'. That is, the intuition that (justified) belief-fixation is not directly voluntary. Then I go on to introduce habits-endorsement expressivism as a theory that can tackle these problems. At the heart of the theory lies the notion of epistemic habits, namely, habits of justified belief-fixation that with its theoretical versatility and flexibility is meant to help us out of the problems.

Here, three preliminary points should be made clear, though. First, the whole thesis should be seen as an *exploration* of the concept of epistemic justification and, thereby, 'the epistemic justification puzzle' that emerges out of the relevant background assumptions. As such it explores the concept of epistemic justification and the resources and potential of a certain theoretical approach to 'the epistemic justification puzzle'. As an exploration, this is far from claiming to be a conclusive solution to the puzzle. Indeed, given the difficulty and complexity of philosophical problems that makes them persist for thousands of years, it would have been unforgivably naïve to claim anything else[21].

But although I am realistic enough to ascertain that I won't reach the 'Ithaca' of a conclusive solution, I do hope to make an interesting exploration 'full of discoveries'. I do hope to explore and chart some new interesting ground. In the case that the reader does find this ground neither interesting nor new, then I hope that the discussion will at least stir him to think of the problem himself. Perhaps, stir him enough to embark and explore a different, more promising conceptual route to the 'Ithaca' of a conclusive solution to the puzzle. Besides, if I may *meta*philosophize, philosophy is perhaps more about learning through the dangerous journey of exploring different conceptual routes to the Ithacas of conclusive solutions than the ultimate arrival to the elusive Ithacas of conclusive solutions[22].

Second, I have myself serious worries about an expressivist solution to the puzzle, but I still think that *on balance* this is the most promising research project. In

---

[21] Compare Moore (2000): 'Philosophical questions are so difficult, the problems they raise are so complex, that no one can fairly expect, now, any more than in the past, to win more than a very limited assent'.

[22] See B.Russell (2001) for some discussion of the beneficial effects of philosophy.

particular, I am deeply sceptical about the semantic problems of expressivism but, alas, I am even more sceptical about the problems of the alternative approaches. However, philosophy demands a choice among unpalatable options (and perhaps this is one of the things that make it fascinating). Thus, if at some points the reader finds the author to be too confident and complacent about the points argued for, then perhaps the enthusiasm of a novice has taken the most of me. I don't mean to imply that I am all that confident about the views I expound.

Third, it is worth repeating that I largely assume that a solution to the semantic problems of expressivism can, one way or another, be found. But if such a solution cannot be found, as some seem to think, perhaps prematurely, then it seems that by pursuing an expressivist solution to the puzzle I am 'flogging a dead horse'[23]. I am struggling to motivate a philosophical programme that despite being interesting, it ultimately leads to a dead end. Yet, although the semantic problems comprise an issue of utmost importance that endangers the very intelligibility of expressivism, they will be scarcely touched in this thesis. Given the scope of the thesis, we must forgo the chance to explore these issues (like some other relevant issues e.g. sceptical challenges, epistemic internalism/externalism controversy etc.) in substantial depth here.

With the argumentative plan of the thesis outlined, in the next section 1.3, I put forward some more motivation for exploring the possibility of expressivist semantics for epistemic discourse.

1.3 Parallel Lives: Meta-Ethics and Meta-Epistemology

In section 1.1, I briefly adduced some preliminary 'starting point' intuitions that incline me to side with the exploratory option of expressivism. In this section 1.3, I present more of the motivation driving the exploration of an expressivist solution to 'the epistemic justification puzzle'. The motivation consists of the fact that ethics and epistemology are both normative discourses that in recent debates seem to move in a parallel trajectory[24]. The common normative character of the two discourses and the 'ethics-epistemology parallel' fuels the intuition that expressivism could also be

---

[23] I refer here to M.Schroeder (2008, 2009).
[24] For a nice drawing of the ethics-epistemology parallel see L.Zagzebski (1996), H.Field (2000) M.Chrisman (2007), M.Ridge (2007b), T.Cuneo (2007) and K.Kappel (ms).

fruitfully applied from moral to epistemic discourse.

Let me elaborate on this parallel between ethics and epistemology drawn in recent debates. This is the big picture of the narrative I am going to present. Expressivism about moral discourse rose as a theoretical possibility in response to the difficulties that referentialist approaches (naturalistic reductionism/nonnaturalism) to moral discourse encountered. Similarly, expressivism about epistemic discourse arises as a theoretical possibility in response to the broadly analogous difficulties that referentialist approaches to epistemic discourse encounter.

Yet, expressivism about moral discourse not only arrived as a theoretical alternative to referentialist approaches but also had some relative success. Since the publication of C.L.Stevenson's (1937) 'The Emotive Meaning of Ethical Terms' and A.J.Ayer's (1946) *Language, Truth and Logic* and their expounding of a crude expressivism in the form of 'emotivism', expressivist theories have been sprouting up all over the place[25]. Presumably, the relative success of expressivism can be credited to the fact that it seems to explain many of the desiderata that a theory of the semantics of normative discourse must explain and also bears some alluring theoretical virtues.

This relative success of expressivism about moral discourse stirs the suspicion that an also relatively successful expressivism about epistemic discourse might be possible. If an expressivist solution to 'the moral problem' is an open possibility, then there is no reason to ignore the possibility of an expressivist solution to 'the epistemic justification puzzle'[26]. Besides, if I am right of my reading of W.Sellars'(1997) *Empiricism and the Philosophy of Mind*, this propounds exactly such an inferentialist (expressivist) approach to perceptual knowledge discourse.

This intuition is also backed by the expectation of a *unified* semantic theory of normative discourse. Both moral and epistemic discourses are normative domains concerned respectively with the practical 'what one ought to do' and the theoretical 'what one ought to believe' and their common normative character stirs the suspicion that they need the same semantic treatment[27]. Indeed, arguably, this has been the working hypothesis of referentialists since Plato, whom treated both moral and

---

[25] See R.M.Hare (1952), S.Blackburn (1993, 1998), A.Gibbard (1990, 2003), M.Timmons (1999), M.Ridge (2007a) etc.
[26] I allude to M.Smith's (1994) well-known book '*The Moral Problem*'.
[27] See also L.Zagzebski (1996), T.Cuneo (2007) and K.Kappel (ms).

epistemic notions in a unified referentialist way. Since both 'the moral problem' and 'the epistemic justification puzzle' concern normative issues and an expressivist approach to moral discourse appears possible, then an expressivist approach to epistemic discourse should also appear possible.

Actually, this is the intuition that inclines me to oscillate sometimes from talk of the semantics of normative discourse to talk of the semantics of epistemic justification discourse. I am inclined to think that the semantic treatment of normative discourse should be the same for both moral and epistemic discourse and stand or fall together[28]. Let us now present a more detailed narrative of 'the ethics-epistemology parallel' in contemporary debates[29].

The parallel can be crisply presented in terms of a question: *How should we understand normative (moral and epistemic) concepts*? In recent debates, attempted answers to this question have led to the formulation of a dilemma of theoretical approaches[30]. The first horn of the dilemma is *Analytic Naturalistic Reductionism[31]*. It suggests that normative concepts can and should be reductively analysed in terms of non-normative, natural properties. Following the analytic naturalistic reductionism project, it has been suggested that moral concepts like 'goodness' and 'justice' should be respectively analysed as 'pleasure' and 'giving back what isn't yours'[32]; or epistemic concepts like 'justification', 'knowledge', and 'truth' should be respectively analysed in terms of 'reliable belief-forming processes', 'true justified belief', and 'the correspondence relation between truth-value bearers and truthmaking or falsemaking facts'[33].

Unfortunately, such attempts seem to meet Moorean 'open question' semantic intuitions that allow counterexamples to emerge and defeat the suggested definitions. Moore (2000) famously argued that is always possible for a clear-headed agent to

---

[28] Compare J.Kim (2008:548): ' In any event, once the normativity of epistemology is clearly taken note of, it is no surprise that epistemology and normative ethics share the same metaphilosophical fate'.

[29] The presentation of 'the ethics-epistemology parallel' draws from my 'Naturalism and Normativity' (2008).

[30] The dilemma is a very old one going all the way back to Plato. See my (2009) 'Plato, Moore and Ethical Nonnaturalism'

[31] I ignore here the recently arrived sophisticated synthetic naturalistic reductionism.

[32] The analysis of goodness as pleasure has been suggested by Plato in *Protagoras*, J.S.Mill (1991) in *Utilitarianism* and in various ways by modern Utilitarians. The analysis of justice as 'giving back what isn't yours' appears in Plato's *Republic I* and is subsequently refuted by the Socratic character.

[33] The reliabilist analysis of justification has been prominently proposed by A.Goldman (1991) and the analysis of truth as a correspondence relation, among others, by B.Russell (2001).

question, without any semantic confusion, whether a moral property M (e.g. goodness) is reducible to a natural property N (e.g. pleasure). In short, the reduction of a moral property to a natural property remains always an 'open question'. Moore assumed that if there was such an a priori property identity, a clear-headed agent wouldn't fail to grasp it and carried on- from the fact that the semantic 'openness' founders all such reductive efforts- to draw the conclusion that moral properties cannot be reduced to naturalistic terms. He contended that all such reductive attempts have committed -what he called- 'the naturalistic fallacy' because there are no natural properties that could reduce moral properties.

Analogous Moorean 'open question' semantic intuitions seem to apply to virtually all core moral and epistemic concepts. These 'open question' semantic intuitions undermine attempts to reduce such normative concepts to naturalistic terms and allow counterexamples to pop up and refute the proposed reductions. Thus, intuitively, there are respective instances of bad pleasures e.g. sadistic cat killings, there are cases where the just thing to do is not to give back what isn't yours e.g. when someone lends you a weapon and then she goes mad, there are E.Gettier (1963) cases of epistemic luck where true justified beliefs are not instances of knowledge e.g. the classic cases with sheep-like dogs, façade barns etc., cases where beliefs formed by reliable cognitive processes are not justified e.g. Euclidean geometry in the light of Einsteinian physics, cases where there are no natural truth-making facts of seemingly true beliefs e.g. mathematical, logical facts etc.

What all these examples show is that such normative concepts seem to resist successful reductive analysis and this decisively undermines the project of analytic naturalistic reductionism. It appears then, that there are no natural properties that can reduce such normative notions. It should be noted, though, that Moorean 'open question' considerations are, as W.Frankena (1939) pointed out, question-begging but still such considerations can be taken to be *enthymatic[34]*.

As Frankena rightly objected, Moore's 'open question argument' begs the question against the analytic naturalistic reductionist because it assumes that there is no such reductive analysis of moral concepts to be found. This seems to trivialize the 'open question argument' because it might be the case that we just didn't *yet*

---

[34] See also B.Williams (1993).

discover the right analysis of moral (and other normative) notions and such reductive analyses will be discovered sometime in the future[35].

However, many anti-reductionist philosophers (nonnaturalists and expressivists alike) think that there is something valuable in the argument that should not be lost by the, indeed, correct Frankena objection. The argument indicates that, intuitively, our constant failure to find such successful reductive analyses inspires a legitimate pessimism about the prospects of such a discovery. Actually, the anti-reductionist pessimism is so pervasive that, even if we find a reductive analysis that seems to be initially immune to counterexamples, what we will think is that we haven't yet found the right counterexample, not that we have found the right analysis[36]. Despite the fact then that the argument is question-begging and inconclusive against analytic naturalistic reductionism, it need not be taken to be trivial. It can be considered to be an *inference to the best explanation* of our constant failure to reach reductive naturalistic analyses of normative notions[37].

The constant failure to provide successful analyses of normative concepts has frustrated many philosophers who accepted the verdict of Moorean 'open question' considerations and turned to the second -Platonic- horn of the dilemma: *Nonnaturalistic Anti-reductionism.* Nonnaturalistic Anti-reductionism accepts that normative concepts are unanalysable concepts that purport to refer to the corresponding nonnatural properties (goodness, justification etc.). Examples of nonnaturalists abound. Plato in his middle period dialogues (*Phaedo, Republic, Theatetus*), arguably, took the properties of goodness, justice and knowledge to be irreducible and nonnatural (or in Platonic jargon, 'forms') and G.E.Moore (2000) also famously proclaimed that the property of goodness is 'simple, indefinable, and sui generis' and any effort to reduce it in naturalistic terms commits 'the naturalistic fallacy'.

Finally, T.Williamson (2000) has also argued that the concept of knowledge is basic and unanalysable. Further, philosophers like H.Putnam (1978:107-9) and E.Sosa (2001) have also directly appealed to Moorean 'open question' considerations

---

[35] This has led J.Searle (1969) to dub it 'the naturalistic fallacy fallacy'.

[36] Compare Williamson (2000: 30-1): 'Even if some sufficiently complex analysis never succumbed to counterexamples, that would not entail that identity of the analysing concept with the concept *knows*. Indeed, the equation of the concepts might well lead to more puzzlement rather than less'.

[37] For such refinements of the classical open question argument see M.Ridge (2003) and A.Miller (2005).

in order to claim that the property of truth is unanalysable while D.Davidson (2001a) has appealed to analogous Moorean considerations in the Platonic corpus in order to reach the same anti-reductionist conclusion. They all conclude that the property of truth is irreducible and nonnatural.

Unfortunately again, the nonnaturalist horn meets some serious epistemological and ontological difficulties of its own that undercut its prospects of success. Here is the rough form of two of them. First, there is 'the epistemic access problem'. Nonnaturalists often posit a cognitive faculty of rational intuition that supposedly can track normative nonnatural properties through some sort of 'intellectual seeing'. But given that we have epistemic access to natural properties, objects, events etc. with ordinary causal-perceptual processes, it appears mysterious how we can have epistemic access to acausal, nonnatural normative properties. Thus, talk of 'intellectual seeing' appears epistemologically spooky, unless there is a way to cash out the ocular metaphor in naturalistic terms.

Second, there is 'the queerness problem'. It is hardly clear what such acausal nonnatural properties could be and where these properties could be placed in the spatio-temporal, natural world. For surely they must be radically different from the mundane natural properties, objects etc. we perceive in our everyday life. If there were such properties, in J.Mackie's (1977) often-quoted words, they would seem to be ontologically 'queer'. They would be radically different from ordinary natural properties like being water, honey, chair, tree etc.

The respective difficulties of the naturalistic reductionism and nonnaturalistic anti-reductionism horns of the dilemma have ushered some philosophers to search for a different approach to normative concepts. That is, a different approach that could avoid at least many of the difficulties that surround and undermine the two horns. One approach that was born from this endeavour (in C.L. Stevenson's (1937) and A.J.Ayer's (1946) hands) was expressivism. Stevenson and Ayer was, of course, applying expressivism only to moral discourse (though Ayer's (1946) treatment of the a priori is, I suspect, an expressivist account in germ form). But contemporary expressivists like S.Blackburn (1996) AGibbard (2003) H.Field (2000, 2009) M.Ridge (2007b), M.Chrisman (2007) and K.Kappel (2010) have also applied

expressivism to epistemic discourse as well[38].

Expressivism, like nonnaturalism, accepts the verdict of Moorean 'open question' considerations and takes such normative concepts to be irreducible to naturalistic terms. Crucially, however, it denies 'the referential semantics assumption' that both horns of the dilemma share. Namely, it denies that there are normative properties and, since our perceptual mechanisms do not indicate the existence of such normative properties (natural or nonnatural), it comports nicely with a parsimonious naturalistic framework. For the expressivist approach, there are no normative properties (natural or nonnatural) that normative concepts purport to refer to. Actually, expressivist semantics are not referential at all. They are use-theoretic semantics explaining the meaning of normative sentences in terms of how the sentences are being used.

Nevertheless, in spite of the rejection of 'the referential semantics assumption', sophisticated expressivists like, S.Blackburn (1984,1993,1998), M.Timmons (1999), A.Gibbard (2003) M.Ridge (2007a) and M.Chrisman (2007) in various different ways claim that expressivism can have 'the best of both worlds'. That is, can still latch onto normativity and reconcile it with such a parsimonious naturalistic framework. They claim that expressivism can mimic the realist-seemings of normative discourse and adequately explain them without any loss of explanatory power. To that effect, they propose various ways in which this could be accomplished. Quasi-realists like Blackburn (1996) and Gibbard (2003) endorse a *minimalist* 'ontologically light' reading of notions like 'property' 'truth' 'fact' etc. and claim that expressivism can legitimately 'earn the right' to talk about minimalist normative properties, truth etc. 'as if' these existed, without conceding that traditional realism can do any better[39].

By denying 'the referential semantics assumption' of both naturalistic reductionism and nonnaturalistic anti-reductionism, expressivism draws out the rug from under the feet of both horns of the dilemma and opens some interesting logical space between the horns of the dilemma. This new logical space promises to elude the difficulties of naturalistic reductionism and nonnaturalistic anti-reductionism.

---

[38] And it has been argued that it reaps some interesting explanatory fruit as well. Kappel (2010) argues that with an expressivist approach to knowledge discourse we can solve 'the Meno problem', namely, the problem asking why knowledge is more valuable than mere true belief.
[39] For the classic exposition of deflationism about truth see P.Horwich (1990).

On the one hand, expressivism comports nicely with a broadly naturalistic framework, as naturalistic reductionism would like it to be, while at the same time explains our 'open question' semantic intuitions concerning our efforts to analyse normative concepts. For expressivists, 'open question' semantic intuitions indicate that there are no natural properties that can successfully reduce our normative concepts. On the other hand, sophisticated expressivists like A.Gibbard, S.Blackburn and M.Ridge, M.Chrisman attempt to show that we can uphold the realist-seemings of normative discourse in an expressivist context, as the traditional realist nonnaturalist would like it to be, while at the same time remain committed to a broadly naturalistic framework that evades -among others- 'the epistemic access' and 'queerness' problems of nonnaturalism.

In conclusion, what 'the ethics-epistemology parallel' indicates is that expressivism about epistemic discourse, as expressivism about moral discourse, comes to the fore as a response to the problems of analytic naturalistic reductionism and nonnaturalism. Accordingly, motivated by the relative success of the application of expressivism to moral discourse, I aspire to explore an application of expressivism to epistemic justification discourse that could have at least the same relative success. Prima facie at least, the common normative character of the moral and epistemic discourses calls for the same unified semantic treatment and the parallel drawn between ethics and epistemology in recent debates spurs this intuition.

With the presentation of 'the ethics-epistemology parallel' and how it confers some more motivation for the pursuit of epistemic expressivism explained, in the next section I make some clarifications about several issues related to the epistemic justification puzzle.

1.4 Some Important Clarifications

The epistemic justification puzzle touches a number of issues (some more important, some less important) that we should better clarify a bit before we set sail for our conceptual exploration. These clarifications concern the formulation of the puzzle and other issues related to the notion of epistemic justification.

First of all, I adopt the more natural way of talking and use only the term 'justification' instead of 'epistemic justification' but, of course, mean epistemic justification. Other species of justification like pragmatic, prudential etc. justification

is something that does not interest us and remains entirely off the picture of the thesis. However, a short note on the epistemic/pragmatic justification distinction is needed for reasons of demarcation. With epistemic justification I mean the kind of justification that aims at truth while with pragmatic justification I mean the kind of justification that does not aim at truth but at personal benefit or advantage. A good example of a case where pragmatic justification is meant to be in play is Pascal's Wager (though the exact details of the argument need not concern us here)[40].

Pascal's wager invites epistemic agents to believe in the existence of God quite independently of whether they think such belief is epistemically justified or not. Rather, they are asked to believe in the existence of God purely on pragmatic grounds: maximum expected utility. Other examples of pragmatically justified beliefs are convenient self-deception cases where the agent holds epistemically unjustified beliefs about himself, but these beliefs are pragmatically justified because they seem to benefit him in some way (e.g. enhance his social or athletic performance).

Second, someone might think that the formulation of the puzzle is still somewhat strained in terms of ordinary language. We don't seem to often use sentences like 'S justifiedly believes that p' or 'My belief that p is justified' in ordinary linguistic practice. Epistemic justification is 'a philosopher's concept' and it is not all that frequently used in ordinary epistemic discourse. But this should not be seen as a problem because, although it is more of a philosopher's concept, it stands out for many mundane epistemic expressions of laymen like 'S's belief is well-grounded', 'p is well-supported', 'p is evidence-based' etc. It is just that philosopher's have an umbrella term of art that covers all these epistemic expressions. This leads to the next point.

Third, the formulation of the puzzle is in terms of epistemic justification but the same positive proposal, namely, habits-endorsement expressivist approach is meant to account for the semantics of other epistemic sentences conferring positive epistemic status. Such other epistemic sentences are 'S's belief is well-grounded' or 'S's belief is supported by good or sufficient reasons' or ' S's belief is reasonable or rational' etc. For reasons of simplicity and clarity, though, I will be talking only in

---

[40] For the argument itself see A.Hajek (2004).

terms of justification.

Fourth, critics might complain that my thesis – as outlined in section 1.3- totally ignores two well-known approaches to normative issues: error (or fictionalist) and relativist approaches. Both approaches accept all four premises of the puzzle and therefore there is not really a philosophical puzzle. It has the appearance of a puzzle but is rather a straw man, they might say.

Let us first take up the error approach. Critics might claim that I am helping myself because from the criticism of reductionist and nonnaturalist approaches I jump to the conclusion that 'the referential semantics assumption' is plausibly mistaken. But this jump is unwarranted (or at least too quick) because there is a referentialist approach, namely, error theory that claims that although there are no such normative properties, we nonetheless purport to refer to such properties[41]. Thus, referential semantics are actually consistent with the claim that there is no natural or nonnatural epistemic justification property and, therefore, we need an argument against error theories before we conclude that 'the referential semantics assumption' should better be rejected.

This is true as far as it goes. I don't discuss the possibility of an error theory approach to epistemic justification. One reason that I ignore 'error theory' is because I don't really know anyone propounding the view about epistemic justification, though, the logical space for such view is wide open (and has been applied to at least moral, modal and mathematical discourses). A second reason is the scope of the thesis that is already quite extended. The breadth of the topic I am addressing is vast and for this reason I have to be a bit selective of the positions I examine. This second reason is related to a third. I consider the error theoretic approach to be one of the least attractive and given that I need to be selective I chose unashamedly to ignore this position. This might seem unfair but I am afraid is the plain truth.

Two more reasons in support of the view that error-theory approaches are one of the least attractive are the following. First, error theories seem to explain away objectivity too easily and this, I think, is a liability. Any theory of normative discourse is expected to accommodate certain desidarata, rather than try to sidestep them by providing a debunking explanation that simply refuses to take them

---

[41] Error theories about moral discourse have been defended by J.Mackie (1977) and R.Joyce (2007).

seriously. One of those desiderata of normative discourse is surely objectivity and error theories seem to give a debunking explanation to this desideratum.

Of course, raising this worry against error theorists might surprise some philosophers because my favoured expressivist approach to 'the epistemic justification puzzle' is also anti-realist and therefore they might presume that it meets the same problem. This is I think too quick. True enough, the standard picture is that precious truth and objectivity can be safeguarded only within a traditional realist framework. Yet, this picture has been contested by expressivists of various kinds who although remain staunch antirealists don't relinquish truth and objectivity[42]. Indeed, this is what I think expressivism's great promise is as a research programme: to account for the realist-seemings of normative discourse in an antirealist framework. I have already referred to such expressivists that want to have it both ways in the last section 1.3[43].

Second, although error theorists do reap some important explanatory fruit (e.g. dodge the ontological and meta-epistemological burden etc.) their way of thinking does not go deep enough because it still remains referentialist in character. By my lights, this is mistaken because the root of the problem lies in our taking the referentialist pretensions of normative language too seriously and this way of thinking should be abandoned completely. Error theory only goes half the way and retains relics of 'the referentialist paradigm'. Obviously, this is way too brief and entirely inadequate to dispose of error theory, but I can't here go any deeper into the substance of the approach. On this point, I am afraid that all I can say is that I will assume the falsity of error theory approaches and cite the work of other philosophers who have argued against the possibility of an error theory approach[44].

Let us now turn to relativism. It is true again that I ignore a relativist approach to the epistemic justification puzzle and people with relativistic leanings might find this partiality disappointing. Perhaps, they will find this partiality especially disappointing because they are in agreement with the view I intend to argue for, namely, that epistemic justification is neither reducible to a natural property or an

---

[42] This attitude has also raised the question of how then we can tell the difference between traditional realists and expressivists realists. J.Dreier (2004) calls this 'the problem of creeping minimalism'. For some discussion see M.Chrisman (2008).

[43] See also M.Chrisman (2008) on the realism/antirealism distinction.

[44] See J. McDowell (1998: 131-151), S.Blackburn (1993 : 149-166) and M.Timmons (1999).

irreducible sui generis property.

Relativists typically agree that, one the one hand, there is no justification property to be reduced and that, on the other hand, there is no sui generis, nonnatural justification property[45]. They then proceed to propound relativist theories of various sorts where justification assertions and attributions are taken to be *relative to* a single person, community, society, and in extreme cases perhaps even gender, social class, nation, race etc. But what unequivocally unites these various relativist theories is the shared feature that justification assertions and attributions are relative to a certain variable, no matter which variable is this. Their entrenched intuition is that justification assertions and attributions cannot somehow be disengaged from a certain point of perspective, context or variable as the realist dreams. There can be no Archimedean point of view.

Different relativists would solve or dissolve the epistemic justification puzzle in different ways. S.Stich (1998) would, for instance, I think, accepts all four premises of the puzzle and claim that relativism defuses any tension between the premises of the puzzle. He would accept that we purport to refer to a certain property of epistemic justification (P2) but that actually there is none, natural or nonnatural (P2 and P4), and that we ultimately have to confront and accept 'the problem of cognitive diversity', if we are not to be 'epistemic xenophobes', as an inextricable fact of social life. Rorty (1979) would probably deny the very existence of such a puzzle from the start and would contend that there is a puzzle if and only if we assume the referentialist 'language game' of analytic philosophy but, of course, we need not do this. It is only when we accept the referentialist 'language game' of analytic philosophy that the puzzle forces its way to the scene but once we become self-conscious and renounce this language game the puzzle simply dissolves[46].

Nevertheless, the reason I chose to sideline relativistic approaches is because again they seem to explain away our truth and objectivity intuitions too easily. Like error-theory, they don't seem able to account for the truth and objectivity intuitions that our epistemic claims seem to imply. And as I also said above, it has been argued that these realist intuitions need not be tied to reference as standard referential

---

[45] See for S.Stich (1998) and R.Rorty (1979).
[46] There are also expressivist relativistic approaches. H.Field (2000, 2009) for example is happy to propound a form of expressivism about epistemic justification justification that is 'moderately relativistic'.

semantics assume. Nonreferential semantics could also claim truth and objectivity intuitions for normative sentences; or at least this is what they argue for.

Fifth, the notion of epistemic justification lies at the heart of the epistemic internalism/externalism controversy and whatever we might say about epistemic justification, inescapably, will have some bearing on this controversy. Let us briefly see what the controversy is about and then what our stance on it is. There are different ways to draw the line between internalism and externalism but for present purposes we can ignore such complications and understand it in pretty much a standard way[47]. Simplifying somewhat, the internalist claims that for an epistemic agent to have a justified belief it is necessary to have access to the good grounds or basis for that belief (e.g. L.Bonjour (1985)). Call this internalist claim *'the accessibility requirement'*.

The externalist contests 'the accessibility requirement' and denies that access to good grounds or basis is necessary for a justified belief. He points out that we can have justified beliefs formed by reliable processes or faculties, though the grounds of these beliefs are beyond our conscious grasp (e.g. Goldman (1991)). The internalist's rejoinder then is to resist the claim that such beliefs should count as genuine instances of justified belief.

Accordingly, internalists think that epistemic justification is a necessary condition for knowledge and externalists counterclaim that justification need not be a necessary condition for knowledge, if knowledge is based on reliable belief-forming processes (Goldman (1991)). In other occasions, externalists might accept that epistemic justification is necessary for knowledge but clarify that epistemic justification should be understood not in a 'Cartesian' way but in an 'anti-Cartesian' externalist way. That is, in the 'anti-Cartesian' terms of reliable belief-forming processes where the epistemic agent is not in a 'Cartesian' way expected to have access to the grounds or basis of the belief.

My stance on the controversy is reconciliatory. Like other philosophers that have a reconciliatory approach to this perplexing controversy (Blackburn (1993:33-52), Brandom 2000, Alston 2005 etc.), I think the best way to go forward is to compromise the two valuable intuitions that respectively internalist and externalist

---

[47] See G.Pappas (2005).

positions emphasize. The internalist emphasizes that is epistemically good to have access and be able to provide reasons or grounds for your beliefs. The externalist emphasizes that often we seem to hold justified beliefs for which we don't have access and can't provide good reasons or grounds.

The two intuitions at work in both sides of the controversy are important and can be reconciled in a 'mixed' approach aspiring to transcend the controversy. This is, of course, easier said than done but I think is a plausible approach and the habits-endorsement expressivist proposal will tend to presuppose such a 'mixed' stance on the controversy. For present purposes I think that is sufficient.

Sixth, a short note is needed on why I am pursuing an account of the nature of the notion of epistemic justification and not knowledge itself, the golden but elusive dream of every epistemologist. The motivation stems from my sympathy to the idea that we should better start with the weaker notion of epistemic justification and from there proceed to establish, if we can, a weak relation to knowledge itself rather than conversely. This view is not novel in the literature, though, of course as much else is contentious[48]. Here is C.Wright (1991:88) explaining such a stance: 'knowledge is not the proper central concern of epistemologico-sceptical enquiry… We can live with the concession that we do not, strictly, *know* some of the things we believed ourselves to know, provided that we can retain the thought that we are fully justified in accepting them' (Wright's emphasis)[49].

Unlike other approaches that conversely take the notion of knowledge as basic, 'the unexplained explainer' we must postulate in order to explain the rest epistemic phenomena (e.g. T.Williamson 2000), I am inclined to think that we should start from the weaker notion of epistemic justification and from there cautiously try to build bridges to knowledge. The order of explanation that starts from the assumption that knowledge is basic seems to my eyes to take thinks backwards, to put the cart before the horse as one might say.

One of the difficulties for 'the knowledge first view' is that in Moorean fashion takes knowledge as something 'given', as a foundation form where we can proceed to understand the rest epistemic phenomena (belief, justification, evidence etc). But

---

[48] Compare J.Pollock (1985:9) 'The central topic of epistemology is epistemic justification rather than knowledge' and L.Bonjour (1985: 5) : 'The concept of epistemic justification is clearly the central concept in the whole theory of knowledge…'. See also R.Fumerton (2002).

[49] Wright calls this stance 'the Russelian retreat' because it is advocated in B.Russell (2001).

surely this will raise the eyebrows of people with sympathy for sceptical challenges. For such philosophers hold that knowledge is something that needs to be argued for, not just to be assumed as a 'given'. Thus, the existence of knowledge is a very controversial and important premise to be conceded 'on the cheap'.

In contrast, one of the advantages of taking the 'epistemic justification first view' is that it exactly gives full heeding to sceptical challenges about knowledge and worries and from there attempts to build bridges with the valuable notion of knowledge. In this way, I think, knowledge might be established even contra to at least some radical scepticism scenarios. I am quite sympathetic to self-defeat and transcendental argument responses to scepticism that try to stress that the sceptic in order to even articulate his sceptical challenge in the first place needs to presuppose at least some knowledge[50]. Yet, although I am optimistic that at least *some* sceptical challenges could be defused, once again this task lies outside the topic of the thesis, though, this does not affect the argument of the thesis. For our stance on epistemic justification is compatible with both sceptical and anti-sceptical scenarios about knowledge.

Surely, on the one hand, if we can respond to sceptical challenges and show that at least some of our justified beliefs amount to knowledge that would be beneficent. We would be in a position to show that at least some of our justified beliefs 'hit the target' and amount to knowledge. But, on the other hand, even if we cannot adequately respond to the sceptical challenges, this should not adversely affect our position. For our position remains compatible with sceptical challenges.

We are not defending the inconsistent with sceptical scenarios claim that at least some of our beliefs amount to knowledge. We are only defending (or better we will be defending) the by far weaker claim that at least some of our beliefs are justified. This claim might raise some eyebrows, since I have already committed myself to expressivism that renounces the existence of an epistemic justification property. The contention that we can have justified beliefs without postulating an epistemic justification property is one of the claims of the thesis that will come in Ch.6 and Ch.7.

---

[50] See for example Putnam's (1981) argument against the brain-in-a-vat scenario. A similar argument exists in Kant's (2003) 'The Refutation of Idealism'. Why Putnam/Kant style anti-sceptic argument cannot refute all forms of scepticism see Williamson (2000: ch.7).

If the sceptical challenges cannot be defused and our justified beliefs cannot be shown to amount to knowledge, then at least we do in the best possible way what is up to us. We form beliefs in a responsible and considerate way. Demon world or not, it is what we have some control over and lies firmly within our sphere of responsibility. What is not up to us and lies beyond our control is something that we shouldn't be held responsible for.

Seventh, a short note on the doxastic voluntarism/nonvoluntarism controversy that seems to touch the issue of belief-fixation and epistemic responsibility just canvassed above. The discussion of belief-fixation as something 'firmly lying within our sphere of responsibility' might convey the wrong signals. I don't mean to imply that I am committed to some form of doxastic voluntarism, though I don't want to imply either that I am committed to some form of doxastic nonvoluntarism.

Once again my stance to such sharp distinctions is reconciliatory with a hope of transcending the distinction. Both sides of the divide seem to emphasize a right intuition. Doxastic nonvolutarism emphasizes the intuition that we don't really seem to decide what to believe when we form a belief. Doxastic voluntarism emphasizes the intuition that we don't really seem to be just mechanically forming beliefs either. The two intuitions could be reconciled in a 'mixed' approach that accommodates both stances and also ensures the viability of epistemic responsibility. If we have *some* control on belief-fixation, then epistemic responsibility is not endangered and this is the assumption that we will be working on in the thesis.

Actually, if the hard-line doxastic nonvoluntarist (e.g. W.Alston 2005:Ch.4) is right that there is no sufficiently rich control of our doxastic attitudes to ground epistemic responsibility, then this seems to lead to a quite paradoxical conclusion. Namely, that the normative discipline of epistemology itself is just 'idle talk' because it cannot really make any real difference in terms of our everyday belief-fixation practices. For one of the main aims of epistemology is to prescribe what, if anything, 'one ought to believe'. How we should transform our epistemic lives in order to become better epistemic agents.

Interestingly, people who espouse such a hardline doxastic involuntarism view, go on to talk about epistemic justification (and perhaps knowledge) as if their view has no implications on the field of epistemology. But if they are right that there is no

epistemic responsibility, then they shouldn't be writing treatises on epistemic justification. Properly, since belief-fixation it is not up to us, they should not only be committed to doxastic nonvoluntarism but also scorn debates on epistemic justification and 'what one ought to believe'. We will touch this problem again in Ch.7 where I will suggest that the habits-endorsement expressivist approach to the semantics of justification assertions and attributions could illuminate the debate.

These clarificatory points, admittedly, touch immensely complex issues in an entirely cursory and inadequate manner. Yet the predicament is such that we can't really pursue many of these issues to substantial depth. But in order to avoid leakages (or even shipwreck) before we even leave port for our philosophical exploration, we needed to plug some holes by clarifying some of our commitments. In the next section 1.5, I summarize the work been done in this first chapter and explain how the argument unfolds in the next chapter.

1.5 Conclusion and Summary of the Argument

In this first introductory chapter I have performed four basic things. First, in section 1.1, I introduced 'the epistemic justification puzzle' and explained what motivates each one of the individually plausible but jointly inconsistent premises of the puzzle. Then I explained how I intend to explore an expressivist solution to the puzzle and briefly provided some of the 'starting point' intuitions that incline me to sympathize with this approach. Second, in section 1.2, I outlined the argumentative plan of the thesis, said a few things about the content of each chapter and also presented some clue of the contours of the habits-endorsement expressivist I will deploy. Third, in section 1.3, I provided some more motivation for exploring an expressivist solution to the puzzle by drawing 'the ethics-epistemology parallel' in recent debates. Finally, in section 1.4, I closed the chapter with some clarifications concerning the formulation of the puzzle, the epistemic internalism/externalism distinction, the epistemic/pragmatic justification distinction etc.

In the next chapter 2, I take the first step of my discussion of analytic naturalistic reductionism. I undertake to introduce what the theoretical commitments of analytic naturalistic reductionist are, clarify what conception of epistemic justification we are after, what conditions an adequate reductive analysis of epistemic justification should satisfy and provide a gloss of the natural/nonnatural distinction. Finally, I introduce

what I call 'the Moorean/Humean lesson'. A lesson intended to undermine attempts to reduce by a priori semantic analysis the notion of epistemic justification (and other normative notions). Chs.3-4 will complement the discussion of naturalistic reductionism (analytic and synthetic) and ultimately motivate the plausibility of 'the naturalism-denial assertion' (P3) of the puzzle

Chapter 2 Analytic Naturalistic Reductionism Introduced

2.1 Introduction

This chapter covers largely preparatory ground for the critical examination of analytic naturalistic reductionism, following in Ch.3. It introduces the project of analytic naturalistic reductionism and prepares the ground for its imminent critique. More generally, Chs.2-4 inveigh against the plausibility of -what I call- 'referential naturalism' in order to motivate the plausibility of the premise three of the epistemic justification puzzle. That is, 'the naturalism-denial assertion' (P3): justification assertions and attributions do not purport to refer to a certain natural property of epistemic justification. Analytic naturalistic reductionism is criticized in Ch.3 and its synthetic sibling in Ch.4.

With the term of art *'referential naturalism'*, I designate the realist naturalistic approach that accepts 'the referential semantics assumption' (P2) and based on this assumption explores an account of epistemic justification that either (a) reduces epistemic justification by semantic analysis to a natural property (analytic naturalistic reductionism) or (b) reduces epistemic justification to a natural property in the a posteriori pattern of the Kripke-Putnam reduction of natural kinds and other science-theoretic terms (synthetic naturalistic reductionism)[51].

Referential naturalism should be cautiously demarcated from what we may call *'expressivist naturalism'*. With the term 'expressivist naturalism', I designate the antirealist naturalistic approach that denies 'the referential semantics assumption' (P2), (namely, that justification assertions and attributions purport to refer to a certain property of epistemic justification), opts for an ontologically parsimonious framework that denies the existence of an epistemic justification property and explores an expressivist understanding of the semantics of justification assertions and attributions (and other normative notions). The habits-endorsement expressivist theory of epistemic justification that will be sketched in Ch.7 falls under this label.

With the demarcation of referential from expressivist naturalism at hand, the

---

[51] Error theory is an antirealist, referential naturalistic theory, but given that I have set error theory aside I ignore this approach.

overall plan for this introductory and preparatory chapter on analytic naturalistic reductionism is the following. First, in section 2.2, I do some preparatory work essential for the discussion following in the next chapter. That is, I make explicit what the theoretical commitments of analytic naturalistic reductionism are, what particular conception of epistemic justification we are after, what conditions need to be satisfied for a successful reduction by semantic analysis of the epistemic justification property and, eventually, shortly canvass and suggest a gloss of the vexing natural/nonnatural distinction underlying the discussion for 'naturalistic reduction'.

Second, in section 2.3, as a prelude to the critical discussion in Ch.3, I broach the first argument against analytic naturalistic reductionism. That is, I broach the metaepistemic versions of two classic metaethical arguments, namely, G.E.Moore's (2000) 'open question argument' and D.Hume's (1986) 'is/ought argument' and from these arguments draw –what I coin- 'the Moorean/Humean lesson' whose application, I think, seriously undermines analytic naturalistic reductionist pretensions. The undermining application of 'the Moorean/Humean lesson' comes in the next chapter 3. Let us now turn to the first section.

2.2 Analytic Naturalistic Reductionism Introduced

In this section, we lay out some clarificatory and preparatory work (essential for the ensuing discussion of Ch.3) in the four steps just mentioned above. First, let us try to make the theoretical commitments of analytic naturalistic reductionism explicit. The analytic naturalistic reductionist starts from the implicit adoption of 'the referential semantics assumption' (P2). He accepts that justification assertions and attributions purport to refer to a certain property of epistemic justification.

Unlike error theorists (or 'fictionalists'), though, he also accepts that not only our justification assertions and attributions purport to refer to such an epistemic justification property, but that this property exists. The reductionist thinks that an epistemic justification property exists because assumes that at least some of our allegedly justified assertions and attributions are indeed justified. Call this assertion *'the existential assertion'*. But if there are at least some justified beliefs, he continues, then there must be a property of epistemic justification in virtue of which

these justified beliefs are justified[52]. Call this assertion *'the realist ontological assertion'*.

The analytic naturalistic reductionist then thinks that this property of epistemic justification that, if at least some of our beliefs are to be justified, must somehow be 'out there' could in principle be reduced by a priori semantic analysis to a certain natural property. We can reductively analyse epistemic justification in terms of a natural property, a property that exists 'out there' in nature. In other words, we can reduce (in terms of necessary and sufficient conditions) by semantic analysis the property of epistemic justification to a natural property constitutive of its essence. Call this assertion *'the naturalistic analysability assertion'*.

To wrap up, these are the theoretical commitments of analytic naturalistic reductionism:

- The Referential Semantics Assumption: 'Justification assertions and attributions purport to refer to a certain epistemic justification property'.

- The Existential Assertion: 'There are at least some justified beliefs'.

- The Realist Ontological Assertion: 'There must be an epistemic justification property in virtue of which justified beliefs are justified'.

- The Naturalistic Analysability Assertion: 'The epistemic justification property (in virtue of which justified beliefs are justified) can, in principle, be reduced by semantic analysis to a certain natural property'.

The aspiration for a naturalistic reduction by semantic analysis of epistemic justification has some prima facie plausibility. We often aspire to provide reductive definitions of many things in scientific practice and sometimes even in everyday life and, hence, it is not at first instance unreasonable to search for an analysis of epistemic justification in terms of necessary and sufficient conditions. For, arguably, an economist can endeavour to define what inflation is, a sociologist what social class is, a psychologist what depression is, a biologist what reproduction is, a physicist what energy is and even a layman what tomato, game or fishing is.

The obvious reason why both scientists and even sometimes laymen search for reductive definitions is that reductive definitions supply us with an informative

---

[52] Compare A.Goldman (1991: 106): 'I do assume that a justified belief gets its status of being justified from some processes or properties that make it justified. In short, there must be some justification-conferring processes or properties.'

account of what the property, event, state of affairs etc. in question is. If the reductive definition is successful, then this will apply invariantly across contexts and help us identify instances of the property, event etc. in question. It will be used as a criterion for identifying instances of the property, event etc. in question[53]. This could be really useful in the case of highly disputed properties, like normative ones, that is not sufficiently clear what they are or what they involve and often cause a lot of disagreement[54].

Of course, as one might guess, reductive definition is not something that can be easily unveiled, especially about abstract and multifarious notions like, for instance, social class[55]. Those with some background in social theory know exactly what I mean, but even those that do not can see what I am driving to. The point I am driving to it is that reductive definitions are often hard to find because counterexamples lurk in the next corner refuting these definitions. And this is not just the case for abstract science-theoretic and philosophical terms. Even in the seemingly simple case of ordinary notions like 'tomato' or 'game' a reductive definition might prove to be much more difficult that one can initially surmise.

For example, this is the definition of what a tomato is found in a dictionary: 'a round red sharp-tasting fruit with a lot of seeds which is eaten cooked or raw as a savoury food'. As a reductive definition this is inadequate because there are tomatoes with a non-round shape like oblong tomatoes. As a result, it is not necessary for being a tomato to be round-shaped. And I am confident that the definition fails in other respects because there are so many different varieties of tomatoes with varying characteristics; not to mention genetic engineering. The moral here is that, as it is sometimes said, philosophers are not lexicographers and dictionaries are not much help in philosophical matters.

But despite the fact that reductive definitions are difficult to be discovered,

---

[53] As various philosophers have noted (L.Bonjour (1985); R.Fogelin (1994)) the whole debate on the nature of epistemic justification could be set in terms of 'the problem of the criterion', as applied to epistemic justification. I have avoided this way of structuring the discussion because it involves scepticism and, as I said in section 1.4, I set aside the scepticism issue.

[54] Analytic naturalistic reductionism was early Plato's project. In the early 'Socratic' dialogues, the Socratic character strives to reduce moral properties, alas, always unsuccessfully. I sketch early Plato's views on knowledge in 'Early Plato on Knowledge' (ms).

[55] Compare the words of a physicist: 'Coming in so many guises, energy is difficult to define. Even now, physicists do not know intrinsically what it is, even though they are expert at describing what it does and how to handle it' J.Baker (2007:20).

searching for a reductive definition is in principle not unreasonable for sometimes we do succeed to provide successful definitions. Physicists, for instance, say that mass is a measure of how much matter an object contains, that is, the sum of the matter of all the atoms of the object in question. Unfortunately, though, like other normative properties epistemic justification is in all evidence not one of those successful cases. For it eventually turns out that the efforts for a reductive definition of epistemic justification are quixotic. No matter how hard we try to find a natural property that reduces epistemic justification, this does not seem to be forthcoming.

We always seem to run into counterexamples that dash our hopeful efforts. And as we already discussed in section 1.2, this seems to be the case for all key normative (moral and epistemic) properties. Although the search for reductive definitions is often hard business, the search for reductive definitions of normative notions proves to be even harder. Normative notions like goodness, truth, knowledge, rationality etc. for some almost enigmatic reason seem to tenaciously resist reductive analysis.

 I say *almost* enigmatic reason because the reason that impedes reductive analysis of normative notions is not enigmatic at all. It is a *semantic* reason. But the proper explication of this semantic reason plaguing analytic reductionist efforts can wait for the next section 2.3. For the time being, we can confirm that epistemic justification stands as no exception to this, I dare say, rule. It is another *irreducible* normative notion.

Let us now take the second step and elucidate what conception of epistemic justification we are after. The notion of epistemic justification is often employed by different epistemologists in different contexts with different meanings and accompanied with different presuppositions. Because 'there are so many notions of justification in the literature that is difficult to identify a single target of dispute,' some philosophers have observed that the concept of epistemic justification 'seems to be in a conceptual muddle' [56].

These philosophers are not overstating, I think. Philosophers often distinguish between objective and subjective justification (J.Pollock 1985), internal and external (E.Sosa 1991), between strong and weak (A.Goldman 1991), egocentric and socially-based (R.Brandom 1995) etc. At the same time, some tie epistemic justification to

---

[56] I quote from L.Zagzebski (1996: 29-31). R.Fumerton (2002) also briefly makes the same observation.

epistemic duty and permissibility (R.Chisholm 1966; H.Vahid 2005), others to rationality and sensitivity to evidence (R.Feldman and E.Conee 2004; R.Wedgwood 2007) and others to epistemic responsibility (R.Foley 2002) etc.

It seems then that epistemic justification can be employed in a wide variety of senses and it is not easy to pull the strings of all these senses together and identify one target concept of epistemic jusitfcation. I don't say that this is impossible, as there might be a way to show that all these ways of speaking about epistemic justification can be pulled together in a nice systematic way that makes sense of this wide variety of talking about epistemic justification. But pursuing such a systematic and unified target concept of epistemic justification would have proven here a time and space-consuming enterprise. For reasons of economy then, I will make a fresh start of my own on the target concept of epistemic justification.

By my lights, the account of epistemic justification we should be after needs to be rooted in our platitudinous understanding of the concept. As such, an account of epistemic justification should satisfy two minimal intuitions surrounding the concept in ordinary usage: (a) truth-conducivity and (b) fallibility[57]. A conception of epistemic justification running against what these two minimal intuitions permit seems to me to be too revisionary to be of any substantial philosophical interest. For any theory of epistemic justification should be continuous with our ordinary understanding and use of the notion and thereby 'save the epistemic phenomena'.

Of course, there might be good independent argument for going revisionary about epistemic justification but I don't see what this argument could be. Thus, I set aside the possibility of a revisionary conception of epistemic justification and remain allied to our platitudinous understanding of epistemic justification. Let me elaborate a bit more on the two identified minimal intuitions.

The first key intuition, truth-conducivity, suggests that epistemic justification is a means to truth[58]. Justified beliefs, theories etc. are more likely to track the truth and get thinks right than beliefs and theories that are unjustified (e.g. depend on wishful thinking, mere fancy, lucky processes like guesses, premonitions and hunches etc.).

---

[57] K.Kappel (2010; ms) makes a nice distinction between knowledge-norms ('k-norms') and justification norms ('j-norms') and explains how these norms differ. My target concept of epistemic justification here is in agreement with his 'j-norms'.
[58] Compare L.Bonjour (1985:7): 'The basic role of justification is that of a *means* to truth, a more directly attainable mediating link between our subjective point and our objective goal'.

Intuitively, it is more likely that you will find your way to Larissa based on a reliable map (and your map reading skills) than by following a hunch or tossing a coin at every juncture.

Many would add that it is an analytic (or 'conceptual') truth that epistemic justification entails truth-conducivity. L.Bonjour (1985:8), for instance, says that epistemic justification is 'essentially' and 'internally' related to truth. That is, it seems to be true in virtue of the meaning of epistemic justification that it is a means to truth. An agent that has acquired the concept of epistemic justification can grasp that justification is conducive to truth acquisition. If he fails to grasp this, then we would be inclined to think that he does not really have acquired the concept, as he does not grasp what the concept minimally implies.

The second key intuition, fallibility, suggests that any justified belief, theory etc. no matter how thoroughly justified might be it could turn out to be false. In ordinary discourse, we take it that although a belief may be thoroughly justified it might ultimately fail to be true. Although we may have a justified belief this might not be sufficient to deliver truth either because it is based on insufficient evidence or circumstantial, misleading evidence.

It should also be made clear that these two intuitions must be satisfied by any proposed theory of epistemic justification, not just analytic naturalistic reductionism. Even if analytic naturalistic reductionism is not the right approach to a theory of epistemic justification (and 'the epistemic justification puzzle'), as we intend to support, any other adequate theory must satisfy the intuitions of truth-conducivity and fallibility. Otherwise, the theory will seem to run against our ordinary understanding of what epistemic justification as a notion involves.

We can now take the third step and clarify the conditions for a successful reduction by semantic analysis of the property of epistemic justification. The phrase 'naturalistic reduction by semantic analysis' implies the identification of a property with a natural one. That is, the semantic analysis of a property to a natural one in terms of necessary and sufficient conditions. If the target property is adequately reduced to a natural property, then the property is nothing 'over and above' the natural property itself.

As any successful reduction by semantic analysis, the proposal must be immune

to counterexamples. If the analysis is to identify the essential property constitutive of the nature of epistemic justification, then surely this analysis should be unassailable. No possible case could be found defeating this analysis. Any belief instantiating the property of the proposed analysis should be epistemically justified, that is, likely to be true. Call this *'the immunity requirement'*.

In addition, the analysis of epistemic justification should be interestingly informative, if it is to have any genuine epistemic value. Many people assume that any such analysis of a normative epistemic property -if it is to be fully interesting and informative- must be non-circular. It must analyse the normative epistemic property of justification to a non-normative, non-epistemic, natural property[59]. Otherwise, the suggested analyses will threaten to be rendered uninformative or if informative only trivially so. For example, we can say that what is (epistemically) justified is evidence-based or that what is rational is reasonable or that what is good is honourable or praiseworthy, but these analyses are rather superficial and only trivially informative (if at all) because they circularly analyse normative terms to other normative terms.

But other philosophers are more sanguine and point out that we can distinguish between *narrow* circular and *wide* circular analyses. While narrow circular analyses are trivial and vacuous and therefore uninteresting and uninformative, wide circular analyses are sometimes both philosophically interesting and informative. A good example for such informative wide circular analysis constitutes the case of colour concepts. Some philosophers like J.McDowell (1994) analyse, let us say red, by saying that for something 'to be red' is for that thing 'to look red to normal perceivers in ideal environmental conditions'[60]. This analysis is obviously circular because the term in question reappears in the proposed analysis but it still seems to be informative. It informs us of what it takes for something 'to be red'.

But though circular this analysis may be 'innocent', they claim, because our account is wide enough to be interesting and informative as an adequate analysis

---

[59] Compare A.Goldman (1991:105) : '…I want a theory of justified belief to specify in non-epistemic terms when a belief is justified' and J.Kim (2008 :532) : '...the criteria of justified belief must be formulated on the basis of descriptive or naturalistic terms alone, without the use of any evaluative or normative ones, whether epistemic or of another kind'. See also R.Chisholm (1966: 11): 'If we are to solve the problem, we must find a definition of *knowledge* that is not patently circular' (Chisholm's emphasis).
[60] See J.McDowell (1994: 29-31).

should be. Besides, as McDowell suggests, colour concepts 'come only as elements in a bundle of concepts that must be acquired together' (McDowell 1994:31) and this seems to imply that they take such concepts to be inter-definable. That is, to be unanalysable in non-circular terms but still to be adequately analysable in wide circular terms. Given that what really matters in analyses is informativeness and that not only non-circular analyses could be interesting and informative but wide-circular analyses as well, any non-circular or wide-circular account is properly informative will do. Call this *'the informativeness requirement'*.

To wrap up, first, these are the intuitions we found that need to be satisfied by any adequate theory of epistemic justification:

- Truth-conducivity Intuition: Any theory of epistemic justification must account for its truth-conducive nature. That is, how epistemic justification is conducive to truth-acquisition.

- Fallibility Intuition: Any theory of epistemic justification must satisfy the fallibility intuition. That is, how epistemic justification does not entail truth (nor knowledge).

Second, these are the conditions we found that need to be satisfied if we are to discover a successful semantic analysis of the property of epistemic justification:

- The Immunity Requirement: 'A successful semantic analysis of the property of epistemic justification (in necessary and sufficient conditions) must be immune to counterexamples'.

- The Informativeness Requirement: 'A successful semantic analysis of the property of epistemic justification (non-circular or wide-circular) must be informative'.

Finally, the difficult question that now confronts us is how to exactly understand the somewhat vague notion of a natural property and overall the natural/nonnatural distinction. This is the fourth and last preparatory step we take in this section. It is a delicate question that has perplexed Moore (2000) himself and still perplexes philosophers ever since. Moore (2000:13) gave (at least) three different accounts of what a natural property is and later on humbly recognised that his 'attempts to define 'natural property' are hopelessly confused'[61]. The question is still perplexing

---

[61] See T.Baldwin's introduction in Moore (2000).

philosophers who provide a wide array of different accounts of how to define a natural property and how, more broadly, to gloss the natural/nonnatural distinction[62]. I will here briefly suggest an account of natural property and attempt a gloss of the natural/nonnatural distinction, drawing from one of Moore's own accounts. The one he later thought to be the correct one[63].

As you might recall from section 1.1, I there very quickly endorsed the popular Shoemaker-Lewis account of a natural property and suggested that a natural property is an empirical property, one with independent causal efficacy. But this definition might not seem very helpful because the paradigmatic candidate analyses of epistemic justification do not seem in a straightforward, undisputed manner to be empirical properties. That is, properties with independent causal efficacy. They are not properties 'out there' in the natural world that impinge and causally interact with our senses and they are not studied by natural science in any straightforward way. Think, for example, of the property of finding a belief (rationally) intuitive or coherent.

Rather, they are mental (or psychological) properties and it is a matter of heated dispute whether mental properties are themselves natural or nonnatural. They are properties instantiated by the cognitive architecture (or the 'mind') of a person and it is not clear whether the 'mind' of a person is reducible to natural properties (like brain states) or not. Mental properties like finding a belief to be coherent or finding a belief to be rational or being a person with reliable belief-forming processes etc. are properties instantiated by the cognitive architecture of a person and, as such, they are not directly empirical or publicly observable (unless someone adheres to some form of obsolete by now Rylean behaviourism)[64].

These mental properties can be seen as natural if and only if we accept a minimally naturalistic picture of the mental[65]. A minimally naturalistic picture of the mental, if it is to remain broadly naturalistic, must at least accept these two central to naturalism claims: 'the causal closure of the natural realm' and 'the supervenience of the mental on the natural'. These are claims, of course, that any picture of the mental must satisfy if it is to have any luck, not just a naturalistic one and accordingly few if

---

[62] For discussion see M. Ridge (2003).
[63] Moore (2000: 13) acknowledges this in his preface to the second edition.
[64] I allude to G.Ryle (1949).
[65] See J.Kim (2000, 2005) and D.Chalmers (1996).

any philosophers of mind explicitly deny or even question these claims[66].

The causal closure of the natural realm suggests that for any natural effect E there is an independent and sufficient natural cause C. It states that every natural event has a certain independent and sufficient natural cause. Nothing from outside the natural realm need causally intervene with happenings within the natural, though, the principle by itself does not logically or metaphysically preclude such nonnatural causal interventions. That would beg the question against dualists who think that there is such causal interaction[67].

All causal closure says is that we need not postulate such nonnatural, causal intervention for explaining natural events. Such postulation of nonnatural causal intervention is shown to be explanatorily gratuitous and, therefore, redundant because there are sufficient natural causes that can do the explanatory job quite well on their own. No invocation of nonnatural causes is required for explaining natural phenomena. The causal closure of the natural underlies empirical scientific inquiry and abandoning this would imply the rejection of empirical scientific inquiry as such[68].

The second claim, the supervenience of the mental on the natural suggests that a mental property (like being in pain or having a belief) necessarily supervenes on a natural property (like C-fibers firing or other brain states), though, the converse does not hold[69]. As philosophers of mind say, mental properties are being *realized on* natural properties. Unlike normative (moral and epistemic) supervenience, this supervenience claim is not an a priori conceptual one. It is rather an a posteriori one bolstered by our empirical findings in neuroscience showing a clear correlation between mental and natural properties. For this reason, the necessity operator is usually construed in nomological and not logical (or conceptual) terms.

The supervenience claim marks the weakest a naturalistic theory of the mental can

---

[66] Even property dualists like emergentists often want to accept these.
[67] See Kim (2005) for how such strong formulation of the causal closure should be avoided because it begs the question against dualists.
[68] Some philosophers, though, with dualist sympathies bite the bullet and reject causal closure. See for example how T.Crane's (2001) emergentism rejects causal closure.
[69] Philosophers of mind speak in terms of physical properties instead of natural properties but I will speak in terms of natural properties in order to avoid any possible complications. There is big controversy of course over whether natural (social, biological, chemical, meteorological, geological etc.) properties are or will ever be reducible to physical properties. See Kim (2005) for some discussion.

get. If we deny the psychonatural supervenience claim, then we deny that a naturalistic picture of the mental is on the right track (or indeed any picture of the mental) because we accept that mental properties are not being instantiated (and are not being ontologically dependent) on natural properties. Thus, we open the road to strong dualist positions. That is, positions considering mental properties to be 'free-floating', distinct and irreducible to natural properties.

Supervenience is, of course, a weaker notion than reduction and reduction is also a weaker notion than elimination and naturalistically minded philosophers of mind sometimes go after such stronger positions. But these complications need not (and cannot) concern us in this context. For the purposes of our discussion here, I assume that such a broadly minimal naturalistic picture of the mental should be on the right track, though, I remain neutral on what exactly naturalistic picture of the mental might be the correct one.

With the conceded assumption of a minimal naturalistic picture of the mental in hand, we now have the tools for a reinterpretation of the natural/nonnatural distinction. In Moore's (2000) footsteps, I reinterpret the notion of natural property in terms of scientific inquiry. I propose that *a natural property* is one that can be studied by the natural sciences and psychology, where psychology is -according to our minimal naturalistic understanding of the mental- understood as a broadly naturalistic discipline. Accordingly, a nonnatural property is one that cannot be studied by the natural sciences and psychology e.g. being an angel. If there are angels, they cannot be studied by natural sciences and psychology.

This interpretation of a natural property is consistent with the initial gloss in terms of independent causal efficacy because these mental properties are themselves considered to be broadly natural properties and, as such, to be causally efficacious. Of course, how to spell out and defend this claim is a complicated and very contentious story in the debates of mental causation. Causal role functionalists like J.Kim (2000, 2005) and D.Chalmers (1996) defend this thesis about intentional properties (though not about phenomenal properties) but this is something that need not (and cannot) directly concern us here. The conundrum of mental causation is one that, admittedly, we quickly have to sidestep.

This fourth step signals the end of this preparatory section. With the explication

of, first, the theoretical commitments of analytic naturalistic reductionism, second, the conception of epistemic justification we are after, third, the necessary conditions for a successful reduction by semantic analysis of epistemic justification and, fourth, the gloss of the natural/nonnatural distinction in hand, we can go on to introduce the two classic metaethical arguments and apply them to metaepistemic terrain: Moore's 'open question argument' and Hume's 'is/ought argument'. I think they create an interesting case against analytic naturalistic reductionism. The interesting case against naturalistic reductionism will, I hope, become evident from the theoretical introduction of the next section, although the practical application proper to analytic reductionist approaches will come in the next chapter 3.

2.3 The Moorean/Humean Lesson

The preparatory work being completed in the last section, the way opens for the argument against the plausibility of analytic naturalistic reductionism; the argument that seems to support the conclusion that analytic naturalistic reductionism is really a quixotic project.

In this section, I introduce the argument that stems from two well-known classic arguments: Moore's 'open question argument' and Hume's 'is/ought argument'. I then elicit from these classic arguments what I call 'the Moorean/Humean lesson'. I go on in the next chapter 3 to apply 'the Moorean/Humean lesson' elicited from these arguments to theories with analytic reductionist aspirations and examine how they cope. As I have already divulged in section 2.1, they won't cope very well; but first to the introduction of the argument itself.

It is common knowledge that the analytic reductionist literature on epistemic justification is virtually vast and, hence, impossible to explore in its utmost limits in the current context; and not just vast, but also deeply complicated because it addresses delicate issues. But we won't be seduced by the temptation to explore this literature to its utmost limits and in substantial depth here. If we, indeed, had opted to explore this literature to its utmost limits and to some detailed substantial depth here, this would surely have drifted us out of our course for a more comprehensive picture in the debates surrounding epistemic justification. In the light of the reasonably limited scope of a thesis, it would seriously vitiate our more comprehensive aspirations.

Hopefully, though, this won't blunt the sharpness of the argument of the thesis because we have a dialectical 'ace up our sleeve'. This ace makes unnecessary the exploration of this vast literature to its utmost limits and to detailed substantial depth without, I think, adversely affecting the sharpness of the argument of the thesis. The dialectical ace up our sleeve is that we can deliver a strike of surgical accuracy that goes directly to the root of the issue and *exposes* the misguided analytic reductionist pretensions of all these positions.

The arguments that 'deliver this strike of surgical accuracy' are the metaepistemic versions of two classical arguments first applied to metaethical discourse: Moore's 'open question argument' and Hume's 'is /ought argument'. These arguments seem to show that the analytic naturalistic reductionist project is rather *unlikely* to ever succeed. In all evidence, it is only a quixotic task that is doomed to endless repetition and constant failure.

Of course, the view that epistemic justification is irreducible and that the project of analytic naturalistic reductionism is quixotic is not new in the literature. There are a number of philosophers that have subjected analytic reductionist approaches to scrutiny and found them to succumb to counterexamples (J.Pollock 1985; W.Alston 2005). From the adduction of counterexamples and motivated by frustration and the pessimistic feeling that no successful reductionist approach looms in the horizon, they concluded that we should dump analytic naturalistic reductionism as a theoretical project.

But I think what, perhaps, is not so new (or at least not so well appreciated) by these anti-reductionist philosophers in the literature is the almost enigmatic reason that epistemic justification (and the other key normative notions), in all evidence, will continue to resist reduction by semantic analysis. The reason, as is hinted in the previous section, is *semantic*, and goes deeper than mere refutation by way of invocation of counterexamples. Actually, it is the reality of this semantic reason that is *the heart of the matter* because it is what gives rise to counterexamples in the first place. This semantic reason is revealed by the two well-known classical arguments: Moore's 'open question argument' and Hume's 'is/ought argument'.

These two classic arguments seem to bring to the open the semantic reason that explains why analytic reductionist approaches are vulnerable to counterexamples and

failure. Once the semantic reason motivating counterexamples is unveiled in this section, in the next chapter we concisely go through some of the main counterexamples and objections that some of the prominent analytic reductionist approaches run into.

But even this enumeration of problems for prominent reductionist approaches is not meant to be exhaustive. Something like that would have required an entire monograph exclusively devoted to this negative portrayal. Rather, it is only meant to make sufficiently clear the semantic reason revealed by Moore's and Hume's classic arguments that leaves analytic reductionist theories exposed, open and vulnerable to counterexamples. But let us now turn to the two arguments that reveal the semantic reason leaving reductionist approaches exposed and vulnerable and explain why normative concepts, in all evidence, resisted, resist and will keep resisting reduction by semantic analysis.

We are already familiar from section 1.3 with Moore's (2000) 'open question argument'. It was first applied to the moral property of goodness and with this argument, Moore sought to dissolve the analytic naturalistic reductionist pretensions of various kinds of naturalists, among others, 19th ce. utilitarians like J.S.Mill, social Darwinists like H.Spencer, hedonists etc. With his exploitation of 'the open question argument', he reached the conclusion that any effort to reduce goodness to a certain naturalistic property committed what he (infelicitously) dubbed *'the naturalistic fallacy'*. There is no such natural property to be found and the quest for it -like the part of moral philosophy that adheres to it- is constantly committing 'the naturalistic fallacy'. That is, fallaciously striving to reduce goodness to a natural property.

Moore's argument appeals to the semantic intuitions of competent users of moral language. It claims that a clear-headed and thoughtful agent, without any semantic confusion, can always raise doubts and resist the reduction of a moral property M (e.g. goodness) to a natural property N (e.g. usefulness). He can always raise doubts and resist the reduction because there is an inherent semantic 'open feel' in any such reduction. In the face of this inherent semantic 'open feel', an agent can always question and resist the alleged reduction with expressions like: 'I understand what desirability or usefulness is, but I don't see how or why desirability or usefulness is good'.

A few centuries earlier than Moore, Hume (1986) with his 'is/ought argument' pointed to the same anti-reductionist conclusion with Moore. In a famous passage in the *Treatise,* he claimed that any moral 'ought' deduced from a certain referential 'is' remains always doubtful and, hence, resistible[70]. Given that one acts only under 'the guise of the good', no matter whether something is pleasant, desirable, useful, reputable, profitable etc. there is always an inherent semantic 'open feel' between this referential 'is' and the normative 'ought to do'. A clear-headed and thoughtful agent can always question and resist a normative injunction of the form 'x is pleasant, therefore, you ought to x' because stumbles on an inherent semantic 'open feel'.

It is this inherent semantic 'open feel' ubiquitously accompanying efforts to reduce normative properties to a natural property or deduce a normative injunction from a natural property that explains the constant failure of such reductionist approaches. This is the semantic reason that we had been talking about. This *inherent semantic 'open feel'* impedes a clear-headed and thoughtful agent from any quick reduction of a normative property to a natural property or any quick deduction of a normative injunction (an 'ought to do' or 'ought to believe') from a natural property[71].

But we should be careful not to exaggerate the exact logical import of these arguments. It should be made clear that, as we already noted in section 1.3 in relation with Moore's argument, these classic arguments do not provide a *conclusive* case against analytic naturalistic reductionism. Although Moore was impressed enough by his argument to go on and infelicitously call any attempt for analytic naturalistic reduction of goodness as committing 'the naturalistic fallacy', there is no logical fallacy involved in such reductionist attempts.

Analogously, Hume's 'is/ought argument' is also sometimes referred to as 'Hume's Law' but this is equally infelicitous. In any case, it is infelicitous if we take the meaning of the notion of 'law' to be nomological. In the discussion following below, I discuss only the exact logical import of Moore's 'open question argument'

---

[70] See Hume (1986: 521). The argument appears in Book III, Part I, Section I.
[71] Similar attitudes are sometimes expressed in literature. Compare F.Dostoyevsky (2003: 813): 'What kind of belief is it that is forced upon a man? What is more, in the matter of belief no proof is of any avail, especially the material sort. Thomas believed not because he saw the risen Christ, but because he already desired to believe'.

but the same points evidently can be transposed and applied to the equally infelicitous calling of Hume's 'is/ought argument' a 'law'.

Obviously, it is logically possible that there might be a natural property reducing goodness or epistemic justification but we didn't *yet* work hard and ingeniously enough to discover it; at least there is no prima facie incoherence in envisaging such a logical possibility. Moore (2000:5-6) himself later acknowledged that 'it was a pure mistake' to treat the argument as conclusive and took it to be only 'very probably so'. Moreoever, the argument has also nothing to do with the natural in particular either. Moore (2000) himself applied his 'open question argument' to theological-nonnatural analyses of goodness (like divine command theory) and find them failing for the same reasons.

But although Moore was too quick to seize on his argument and infelicitously speak about 'the naturalistic fallacy', his argument does seem to capture something intuitively deep. It captures the fact that the inherent semantic intuitions of 'open feel' constantly undermine reductionist efforts. In the light then of these 'open feel' semantic intuitions and the 2,500 years record of failure in analysing epistemic justification (since Plato's *Theatetus*) it is reasonable to be pessimistic about the prospects of naturalistic reductionism. This pessimism, though, should not be mistaken for something that it is not.

It is not a conclusive rebuttal of reductionism. It can be only seen as an *inference to the best explanation* for our inherent semantic intuitions of 'open feel' and the long history of failures in reducing epistemic justification (and other key normative terms), though, no 'naturalistic fallacy' is involved. Thus, although it is *logically* possible that there is 'out there' in nature a certain property that does reduce epistemic justification and patiently waits to be discovered, in all evidence, it is rather unlikely that the existence of such a property is *metaphysically* possible (as I said, the same considerations apply to Hume's 'is/ought argument').

To end this short digression on the logical import of the classical arguments and revert back to our discussion, the idea is that an epistemic version of these well-known classic arguments can be applied with equal aptitude to epistemic discourse and undermine analytic reductionist approaches to epistemic justification.

The epistemic version of Moore's 'open question argument' claims that a clear-

headed epistemic agent, without any semantic confusion, can always raise doubts and questions about the reduction of an epistemic property E (e.g. epistemic justification, rationality) to any natural property N (e.g. coherence, self-presentation). The reduction will always seem to have an inherent semantic 'open feel' that makes it appear wide open whether the suggested natural property can successfully reduce the epistemic property. Taking advantage of this 'open feel' an epistemic agent can resist the reduction with expressions like 'So what?', 'But this is not really compelling' etc.

It is this inherent semantic 'open feel' that stirs the sceptical suspicion that for any proposed reduction of an epistemic property to a natural property, no matter how elaborate and complex, there is an adequate counterexample just around the corner. The sceptical suspicion surrounding any reductive effort is actually so pervasive that as T.Williamson (2000:31) has nicely put it (in terms of knowledge):

> '[E]ven if some sufficiently complex analysis never succumbed to counterexamples, that would not entail the identity of the analysing concept with the concept *knows*. Indeed, the equation of the concepts might well lead to more puzzlement rather than less'.

Williamson's point, I think, goes deep because it captures the intuition that analytic reductionist approaches to normative notions do not *just* fail because of the emergence of defeating counterexamples; as many antireductionists about epistemic justification in the literature seem to presume. They fail because of the semantic reason we identified, the inherent semantic 'open feel' ubiquitously concomitant of such reductive efforts. Counterexamples are just the inevitable outcome of this inherent semantic 'open feel'. As Williamson says, even if we found an analysis that seems to resist counterexamples, this would 'lead to more puzzlement rather than less' because the inherent semantic 'open feel' prepares us for nothing less than counterexamples.

Of course, in practice, analytic reductionist approaches to normative properties are not so resistant to counterexamples. If they were, then perhaps we would be puzzled -as Williamson foresees - or even tempted to have second thoughts about the irreducibility of normative properties, despite the inherent semantic 'open feel'. But the fact that they are not so resistant to counterexamples saves us the trouble. For instance, suppose we propose that epistemic justification is reducible to coherence.

The inherent semantic 'open feel' of the identification of epistemic justification to coherence will, initially, prevent a thoughtful, though clear-headed, agent from any quick identification of the two. Call this initial stage '*the semantic halt step*'.

Subsequently, the semantic 'open feel' will stir sceptical suspicion that this can't be the right analysis. Once sceptical suspicion intrudes, it is pretty much an exercise of creativity and imagination finding emergent counterexamples that defeat the reduction. This is bread and butter for fastidious philosophers that always relish in defeating theories with exotic (and sometimes plain simple) counterexamples. Someone can claim, for example, that there are cases of coherent beliefs that are not epistemically justified and then proceed to adduce such cases.

Take for example the case of an epistemic agent who thinks that a belief is justified because it coheres and it is supported by his web of beliefs while at the same time overlooks substantial evidence (reliable testimony etc.) against this belief. This is a belief that although coherent, it seems epistemically unjustified and what this shows is that mere coherence is not sufficient for justification. Call this second stage '*the defeating step*'.

The defeating step could also be put in terms of 'multiple realizability'. As in philosophy of mind debates, type-identity reductionism has been considered implausible because mental properties are multibly realizable, that is, realized by the different natural properties of different brain structures[72], analytic naturalistic reductionism is to be considered implausible because normative properties are equally multibly realizable by different natural properties. For example, goodness might supervene in different contexts on varying natural properties like pleasure, usefulness, rational intuitions, desire satisfaction etc. and epistemic justification might supervene in different contexts on coherence, rational intuitions, reliability, epistemic character and virtues etc.

Multible realizability vitiates reductionism because it shows that higher order properties (mental and normative) being realized by different natural properties in different contexts cannot be reduced to any single natural property. Mental and normative properties are being realized by and supervene on many, perhaps infinite, natural properties and this vitiates the possibility of reduction to a single natural

---

[72] For the classic statement and exposition of the multiple realizability argument see H.Putnam (2003).

property[73].

The same conclusion is driven by the epistemic version of Hume's 'is/ought argument'. Given that one ought to believe only what is (best) justified, no matter what natural property a certain belief instantiates we cannot deduce that one ought to endorse that certain belief. A belief might be coherent, rationally intuitive, self-presenting, formed by reliable cognitive processes, motivated by epistemic virtues etc. but it won't be compelling that one ought to endorse that certain belief in question. Let us call this anti-reductionist lesson drawn from the epistemic version of Moore's 'open question argument' and Hume's 'is/ought argument' *'the Moorean/Humean lesson'*.

It seems then that there is a *semantic gap* that needs to be bridged between assigning a natural property to a belief and the *necessary endorsement* of that belief. The 'necessary endorsement' is what this semantic gap of inherent 'open feel' allows a clear-headed and thoughtful agent to question and resist. This is a *crucial* repercussion of all reductionist approaches that assume 'the referential semantics assumption'. Referentialists/reductionists think that the element of necessary endorsement will follow from the correct identification of the property of goodness, rightness etc. (in the moral case) and epistemic justification, rationality etc. (in the epistemic case). But such identification is constantly elusive (if we are right that naturalistic reductionism fails, at least) and, hence, they constantly miss this element of necessary endorsement.

The failure to capture the element of endorsement is important to be recognised and specified because our positive story in Ch.7 will reverse the referentialist order of explanation that seems to take the order of explanation backwards and 'put the cart before the horse'. That is, instead of starting with 'the referential semantics assumption' and seeking to reduce justification in order to capture the element of

---

[73] Two points are in order here. First, someone might suspect that my talk of multiple realizability of mental properties clashes with my commitment to a minimally naturalistic picture of the mental. Yet, this is not the case as multiple realizability clashes only with type-identity reductionism and is consistent with functionalist pictures of the mental. Causal role functionalism as defended by D.Chalmers (1996) and J.Kim (2000, 2005) is a robustly naturalistic picture of the mental –at least as far as intentional mental properties are concerned. Second, as it has happened in the case of mental properties, someone might suspect that epistemic justification is after all, a wildly disjunctive property. That is, a belief if justified if and only if it is either coherent or self-presenting or produced by reliable belief-forming processes or… etc.'. Again, this is too quick, as disjunctive properties do not seem to be explanatorily fruitful. I can't expand on this point here but for some good discussion see Kim (2000: 107-9).

endorsement, it will jettison the referential semantics assumption, opt for a 'use semantic strategy' and start from how we use justification assertions and attributions and the element of endorsement. That is, it will put things right and place 'the horse before the cart'; or so I will argue. Keep this in mind because it is a *crucial* dialectical manoeuvre that will surface again in Chs.6-7.

With the introduction of an epistemic version of Moore's 'open question argument' and Hume's 'is/ought argument' and the drawn anti-reductionist 'Moorean/Humean lesson' under our belt, in the next chapter we turn to the practical application of this theoretical lesson to various approaches with reductionist aspirations. It remains to be seen whether, indeed, theory meets practice, for we know very well that often abstract theory does not square with 'down to earth' practical application.

Let us now take stock before we move on to the practical application task.

2.4 Conclusion and Summary of the Argument

In this preparatory chapter, I have strived to accomplish two things. First, in section 2.2, I made explicit the theoretical commitments of naturalistic reductionism, specified the conception of epistemic justification we are after, clarified the conditions for a successful reduction of epistemic justification and proposed a certain gloss of the natural/nonnatural distinction. This discussion was incumbent on us because these issues underlie the critique of analytic naturalistic reductionism following in the next chapter.

Second, in section 2.3, as a prelude of the ensuing critique of analytic naturalistic reductionism in Ch.3, I introduced an epistemic version of two classical arguments: Moore's 'open question argument' and Hume's 'is/ought argument'. As I argued, these classical arguments explain the semantic reason that efforts to reduce normative concepts (like epistemic justification) systematically founder. The semantic reason is the inherent semantic 'open feel' ubiquitously accompanying any attempted reduction of a normative property to a natural and any deduction of a normative injunction from a natural property.

Based on this semantic reason, I then elicited from these arguments 'the Moorean/Humean lesson', an anti-reductionist lesson intended to be applied with pernicious effect on analytic reductionist aspirations. The 'Moorean/Humean lesson'

suggests that any such analytic reduction of a normative property and any analytic deduction of a normative injunction initially meet 'the semantic halt step' and subsequently 'the defeating step'.

Whether this application of 'the Moorean/Humean lesson' can indeed have this allegedly undermining and pernicious effect on analytic reductionist aspirations, it is something that needs to be decided in the next chapter. Let us turn page to examine whether in this case abstract theory does meet 'down to earth' practice.

Chapter 3 The Problems of Analytic Naturalistic Reductionism

3.1 Introduction

With the preparatory work of Ch.2 in hand, in this chapter we are ready to apply 'the Moorean/Humean lesson' on analytic naturalistic reductionism. The argument has already been introduced in section 2.3 and what remains is simply its practical application. Thus, in sections 3.2-3.3, I step forward to apply 'the Moorean/Humean lesson' -as elicited from the Moorean/Humean classic arguments in section 2.3- against the persistent effort to reduce by semantic analysis the property of epistemic justification.

The practical application of 'the Moorean/Humean lesson' comes in two stages. First, in section 3.2, I apply the lesson to the internalist analytic reductionist approaches of (classical and non-classical) foundationalism, holistic coherentism and L.Zagzebski's (1996) virtue epistemology[74]. Second, in section 3.3, I apply 'the Moorean/Humean lesson' to the externalist reductionist approaches of A. Goldman's (1991) process reliabilism and E.Sosa's (1991, 2007) virtue reliabilism. This application seriously undermines both internalist and externalist analytic naturalistic reductionism. All these efforts are called into question.

Finally, in section 3.4, I close with a conclusion and summary of the argument and prepare the ground for the critique of synthetic naturalistic reductionism following in the next Ch. 4. This conclusion and summary will mark the end of the first part of our critical exploration of referential naturalism.

But an important clarification is incumbent on us before we start examining these approaches. By calling the approaches just mentioned above 'reductive' I have, admittedly, taken some liberties with these theories, as it is not clear that all these

---

[74] I follow the mainstream view and consider the third option of Agrippa's trilemma, infinitism, to be entirely implausible and therefore ignore it. Infinitism is the view that the justification of a belief requires the pursuit of justificatory reasons ad infinitum. One important reason counting against infinitism is that it is cognitively too demanding for cognitively finite beings like us. As Wittgenstein (1953:136) has said: 'Justification comes by experience to an end. If it did not it would not be justification'. Another reason is that it seems to be self-defeating, as it claims that belief in infinitism is justified but this very belief does not pursue justificatory reasons to infinity. But for a heroic defence of infinitism against all odds see P.Klein (2008).

approaches were meant to be reductive by their proponents. I have taken these liberties because the literature explicitly on meta-epistemology is at this stage somewhat sparse, and much of it is on expressivism rather than reductionist analyses. Part of the problem is that meta-epistemology is a relatively nascent field of inquiry and the distinction between first-order, substantive epistemology and second-order, meta-epistemology is not always clearly marked, in a way that is reminiscent of the state of ethics in the early days of meta-ethics[75].

The standard view is that meta-ethics is independent of first-order ethics as someone could be an expressivist about metaethics but a consequentialist or deontologist about first-order ethics. The same view should hold about the epistemic domain, as someone could be an expressivist about meta-epistemology and a coherentist or foundationalist about first-order epistemology. Thus, given the sparse literature on meta-epistemology and that the distinction between metaepistemology and substantive epistemology is often not clearly marked, the characterization of these theories as 'reductive' should be taken with a pinch of salt. I simply speculate about what certain prominent theories would look like if transposed into reductive analyses, without insisting that their actual defenders intend their views to be understood in this way.

With this important clarification drawn, we can now turn to the critical examination of reductionist theories; or at least theories that speculatively could be seen as reductionist.

3.2 The Moorean/Humean Lesson Applied to Internalist Theories

Before starting the examination of internalist positions, three points pertinent to our discussion should be made clear. The first two points have already been touched in section 2.3, but I will reiterate them because they are important. First, this list of theories is far from exhaustive (and it is not meant to be). The literature, for example, only on virtue epistemology is vast and diverse and, reasonably, I can't linger here for too long.

Second, the brief objections cited to each approach are again far from being

---

[75]   Moore's (2000) own reductionist targets about goodness, perhaps, couldn't be classified as 'reductive' (hedonists, utilitarians, social Darwinists etc.) themselves. The distinction in the moral domain still causes disagreements sometimes. See the debate between G.Sayre-McCord (1997b) and M.Smith (1997) on whether M.Smith's (1994) is about 'the moral problem' or 'the metaethical problem'.

exhaustive. Each of the theories I examine is venerable and, thus, deserves a fairly much longer and more attentive treatment than the concise one provided here. Unfortunately, this cannot be done in this inopportune context. This is the reason that in the preface I warned that the thesis intends to capture a more comprehensive picture and, hence, aspires to provide breadth of scope rather than substantial depth.

However, as I explained in section 2.3, I don't think that either the privileging of breadth of scope over substantial depth or the examination of only some paradigmatic reductionist approaches instead of all, adversely affects the sharpness of the argument of the thesis. I think it does not, due to the dialectical ace of 'the Moorean/Humean lesson'. As we have already argued in Ch.2, the 'Moorean/Humean lesson' both strikes with surgical accuracy at the semantic 'Achilles' heel' of these analytic reductionist approaches and, further, in all evidence, it can be used to devastate any other such approach.

Thus, we employ a guerilla 'hit and run' tactic. We hit hard and with surgical accuracy at the semantic Achilles' heel of these prominent theories and disengage quickly. For, otherwise, a full contact engagement with each one of these venerable theories would have proven perniciously time and space-consuming, unavoidably suspending our aspiration for a more comprehensive picture in the debates surrounding epistemic justification.

Third, although these approaches founder as reductive analyses of epistemic justification, this is not to imply that they are valueless from an epistemic point of view. Surely, the notions of coherence, rational intuitions, epistemic virtues, reliable cognitive processes and faculties etc. can sometimes be truth-conducive. But although they may sometimes be truth-conducive, these approaches do not reduce epistemic justification and this is what really matters in this context. They don't satisfy the conditions for a successful reduction by semantic analysis of the property of epistemic justification, as these were made explicit in section 2.2. They don't fill the analytic reductionist's bill, as one might put it.

With these three preliminary points covered, we are now ready for raiding the internalist reductionist camp. Foundationalism is the first internalist analytic reductionist approach to be raided. Foundationalism has been perhaps the most prominent approach in the history of epistemology, cutting across of even the

everlasting rationalism/empiricism dichotomy. Both rationalists like Descartes and empiricists like Hume were, arguably, in some appropriate sense foundationalists[76]. The central idea of foundationalism is that there are certain 'foundational' beliefs that confer justification to other 'non-foundational' beliefs. In what is often called *'classical foundationalism'*, the foundational beliefs are those that in virtue of a specified property are considered to be infallible and the non-foundational are those that in virtue of the absence of that specified property are considered to be fallible.

In virtue of that specified property and (alleged) infallibility, the foundational beliefs are considered as themselves justified and as capable of conferring justification to the fallible non-foundational beliefs. Fallible non-foundational beliefs lacking this specified property are considered justified if and only if they are based on the bedrock of the infallible foundational beliefs. With the foundations/superstructure architecture of justification, foundationalism is often (in philosophical consciousness) figuratively compared with the structure of a pyramid, where the stable foundations are at the bottom of the building supporting the superstructure all the way to the top[77].

An influential defence of classical foundationalism came from R.Chisholm (1966). Drawing from Leibniz, Chisholm (1966:27-9) suggested that the specified property in virtue of which a belief is infallible and, therefore, foundational and justified is 'self-presentation'. As he (1966: 29) says '... what is directly evident to a man is always some state of affairs that 'presents itself to him'. Thus, my believing that Socrates is mortal is a state of affairs that is 'self-presenting' to me'.

Let us measure Chisholm's suggestion against 'the Moorean/Humean lesson'. Is a self-presenting belief justified? (and therefore infallible and foundational). Unfortunately, it seems clear that the account is less than satisfactory. First comes 'the semantic halt step'. A clear-headed and thoughtful agent can appeal to the inherent semantic 'open feel' and halt any quick identification of self-presentation and justification. 'I can understand that p is self-presenting, but I am not quite sure whether it is justified', he might retort. Stirred by these inherent semantic 'open feel' intuitions, he can then take 'the defeating step' and with the help of some creative

---

[76] See E.Sosa (2008) and R.Fumerton (2010) for this categorization. Arguably, Plato and Aristotle were also foundationalists.

[77] See E.Sosa (2008)

imagination adduce counterexamples.

He can contend, for instance, that often beliefs that are 'self-presenting' to the eyes of many people do not really seem to be justified. For example, the belief that 'Stoke city is the best team in Premier League' might be considered 'self-presenting' by many Stoke city fans but most people would think that, although 'self-presenting', it is not really justified. It is not justified because it is based on the feelings of affection for one's favourite team rather than evidence. What this example shows is that self-presentation is not sufficient for justification, as there are ample cases of self-presenting beliefs that are not justified[78].

The same corollary follows from Hume's 'is/ought argument'. A belief may be 'self-presenting' but an agent can legitimately resist inferring that he ought to endorse that certain belief. Just because a certain belief is 'self-presenting', does not necessarily imply that we ought to endorse that belief. For example, for a biased anti-Semite it might surely appear 'self-presenting' that the holocaust is a convenient Zionistic myth. But still, he might be open-minded and honest enough to give a sceptical heeding to the abundant evidence undermining this belief and, hence, resist endorsement of the belief. 'Conspiracy theory' mania aside, there is compelling evidence (survivors, Nazi documents, concentration camps etc.) that the holocaust did occur and even a biased anti-Semite might feel the pressure -stemming from this evidence- to resist endorsement of that belief.

The Moorean/Humean lesson seems to leave classic foundationalism exposed to the emergence of counterexamples. Thus, it does not satisfy 'the immunity requirement' for a successful reduction by semantic analysis of epistemic justification, as this was put forward in section 2.2. That is, it fails to deliver an analysis that yields the essential property constitutive of epistemic justification and, therefore, fails to deliver an analysis immune to counterexamples.

What the inherent semantic 'open feel' of 'the Moorean/Humean lesson' exploits in order to invoke counterexamples and defeat classical foundationalism, is the question of *in virtue of which property* the foundational beliefs are coined foundational and the non-foundational coined non-foundational. The easy answer of

---

[78] Self-presentation can also fail to the other direction too. We can have justified beliefs that are not 'self-presenting' and therefore self-presentation is not even necessary for justification. Thus, self-presentation is neither necessary nor sufficient for epistemic justification. But in any event, insufficiency alone shows that self-presentation cannot reduce epistemic justification.

classic foundationalism is that the foundational/non-foundational dichotomy is respectively symmetric with the infallibility/fallibility dichotomy. He then specifies a property (e.g. self-presentation) that is supposed to identify the infallibility of foundational beliefs. But this can't be a persuasive answer because, as we have seen with Moore's and Hume's arguments, it is not clear at all which property can identify the infallibility of foundational beliefs. Thus, we need the precious property in virtue of which we can discern between infallible and fallible beliefs. Call this *'the discernibility problem'*.

Things get very quickly complicated at this juncture. Some foundationalists think that this discerning property of infallibility is self-presentation, or Cartesian 'clarity and distinctness' of ideas or Humean 'vividness' of ideas while others that is self-justification (or warrant) etc. Yet they all agree that these foundational beliefs are in some intuitive way 'immediately compelling'. The very thought of these beliefs directly inclines the agent to endorse them as true. The fallible non-foundational beliefs are not so difficult to discern. They are the beliefs that are not in some intuitive way 'immediately compelling' themselves and, therefore, their justification must ultimately rest on the bedrock of the foundational beliefs.

But these proposals cannot really blunt Moore's and Hume's arguments. No matter whether we think the discerning property is self-presentation or Cartesian clarity and distinctness etc., Moore's and Hume's arguments show that these accounts are less than cogent. In the same way 'the Moorean/Humean lesson' undermines taking 'self-presentation' to be the precious discerning property of infallibility, it also undermines the rest of the proposals. This is easy to understand, if you have followed the application of 'the Moorean/Humean lesson' on the proposal of 'self-presentation' and grasped how it works.

Thus, discovering which property is the discerning property of infallibility has proven difficult and this has cast some serious doubt on classical foundationalism[79]. The point was quickly seized and exploited by W.Quine (1953, 2008) and W.Sellars (1997) in their famous critiques of (respectively a priori and empirical) foundationalism. The Sellars/Quine influential critique of 'the framework of givenness' has been largely conducive to the waning (at least for some time) of the

---

[79] For discussion see R.Fumerton (2008), E.Sosa (2008) and W.Quine (2008).

popularity of foundationalism.

In the face of 'the discernibility problem' and the Sellars/Quine critique, some philosophers abandoned foundationalism altogether. Others, though, were more sanguine and have attempted to defend a refined form of foundationalism that could avoid the bugbear of 'the discernibility problem'. These more sanguine philosophers thought that we can disjoin foundationalism from the commitment to the problematic infallibility/fallibility dichotomy and, thus, dodge the bugbear of 'the discernibility problem'.

These philosophers concede that is difficult to discern in virtue of which precious property some beliefs are infallible and foundational and claim that we can weaken the theory's commitment in order to make it more flexible and, hence, viable. In order to make the theory more flexible and viable we can reconstitute the infallibility/fallibility dichotomy in terms of *prima facie* infallibility/fallibility. That is, of prima facie infallible foundational beliefs and fallible non-foundational beliefs. This version of foundationalism is often called *'Non-classical foundationalism'*. A recent subtle and influential defender of such a version of non-classical (a priori) foundationalism has been L.Bonjour (1998)[80].

Bonjour (1998:102) suggests that :

'… [a priori] justification… apparently depends on nothing beyond an understanding of the propositional content itself, a proposition whose necessity is apprehended in this way…[it] may be correlatively characterized as *rationally self-evident*: its very content provides, for one who grasps it properly, an immediate accessible reason for thinking that it is true'. Such (a priori) propositions are '…seen, or grasped or apprehended as an act of *rational insight* or *rational intuition*…(1998:102)'.

Let us put Bonjour's (1998) version of non-classical foundationalism to the test of 'the Moorean/Humean lesson'. Is a belief grasped by 'rational insight or intuition' justified? Again, this does not really blunt the classical arguments. A clear-headed agent can appeal to the inherent semantic 'open feel' and resist the identification of 'rational self-evidence' (or grasp by 'rational intuition') and justification. Then he can take 'the defeating step', exploit 'the discernibility problem' and adduce

---

[80] Bonjour (1998) addresses specifically the species a priori epistemic justification but for our present purposes I will assume that his account generalizes and addresses the genus of epistemic justification.

counterexamples.

He can contend, for instance, that often beliefs that are supposedly 'rationally self-evident' in the course of due time (additional reflection, full-information etc.) are shown to be unjustified. Think, for example, of the biased Zionist vehemently denying the brute historical fact that, in order for the state of Israel to be founded after World War II, many native Palestinians suffered forced displacement while being replaced by Israeli settlers from all over the world. For him, this historical fact is 'rationally self-evident' that never took place and, perhaps, this belief has been instigated be anti-Semites. Yet, this belief conflicts with compelling evidence and cannot be considered epistemically justified (though, it is perhaps pragmatically justified).

The same corollary follows from Hume's 'is/ought argument'. A belief may be 'rationally self-evident' but an agent can legitimately resist inferring that he ought to endorse that certain belief. Just because a certain belief is 'rationally self-evident', does not necessarily imply that we ought to endorse that certain belief. Think, for instance, of an agent who thinks that the belief that his wife is not cheating on him is 'rationally self-evident' but still resists endorsement of the belief because he understands that overwhelming evidence suggests otherwise. Thus, like classical foundationalism, non-classical foundationalism also fails to meet 'the immunity requirement'.

Yet, this conclusion is rather unfair to Bonjour's proposal because it neglects Bonjour's commitment to the prima facie infallible/fallible dichotomy and not to the infallible/fallible dichotomy. Bonjour can legitimately respond that 'the Moorean/Humean lesson' should not be a problem for his non-classical foundationalism because it is unashamedly fallibilist and, therefore, can absorb the inherent 'open feel' semantic intuitions and the ensuing counterexamples.

These beliefs are only prima facie infallible and if they eventually fall prey to counterexamples and turn fallible, this does nothing to vitiate his proposal. Counterexamples are consistent with his fallibilist non-classical foundationalism and should not spoil its attractions. Actually, Bonjour (1998) is sensitive to the occurrence of mistaken 'prima facie infallible' beliefs and based on coherence considerations has suggested an account explaining how we can track and correct

false 'rational insights'.

However, the rejoinder from anti-foundationalist epistemologists is not far away. First, taking refuge to a form of refined fallibilist foundationalism seems to lose 'the Cartesian motivation' for being a foundationalist in the first place. For, 'the Cartesian motivation' was to justify our (non-foundational) beliefs by appeal to certain and infallible grounds that can adequately support the edifice of our knowledge. But now we are forced to compromise with something much less promising and ambitious. We are forced to compromise with uncertain and fallible grounds that may not so adequately support the edifice of our knowledge. With this weakening of the commitments of foundationalism, it is not clear whether what is left in foundationalism is what was initially both attractive and desirable.

Second, it seems that non-classical foundationalism *in principle* cannot meet 'the immunity requirement' because denies that it can identify the essential property constitutive of the nature of epistemic justification and, hence, be able to identify justified beliefs. As a result, he accepts the reality of counterexamples by means of fallible prima facie infallible beliefs and, hence, cannot satisfy 'the immunity requirement' in the first place. It cannot because it accepts the inescapability of mistaken prima facie infallible foundational beliefs which are supposed to be themselves justified and also justification-conferring to non-foundational beliefs.

Thus, non-classical foundationalism does not seem to raise high enough to meet the requirements for a successful reduction of epistemic justification. And this is the inconvenient predicament because the foundationalist diagnosing that –due to 'the discernibility problem'- cannot possibly meet the stringent standard of 'the immunity requirement', opts to accept the inescapability of counterexamples by means of fallible prima facie infallible beliefs.

In conclusion, despite the masterful defence of non-classical foundationalism by Bonjour, the thought that, on the one hand, we fail to satisfy 'the immunity requirement' and that, on the other hand, we might ground our edifice of knowledge on possibly shaky grounds is disquieting, and some philosophers have preferred to try out some other approaches to justification. They felt disappointed enough from foundationalism's curtail of aspiration and promise and decided to explore a different and, perhaps, more promising account of justification.

The other prominent traditional approach to epistemic justification has been coherentism. Coherentism has been the option of choice for Hegelian idealists (like F.H.Bradley) but the disillusionment with Hegelianism at the start of 20th ce. and the ensuing rise of early analytic philosophy (through the largely foundationalist work of B.Russell, G.E.Moore and L.Wittgenstein), ushered the popularity of coherentism along the same road with the waning fate of Hegelian idealism. However, since Quine's (1953) attack on analyticity and a priori foundationalism and Sellars' (1997) attack on a posteriori foundationalism, coherentism has not only reappeared on stage, but also enjoyed some rediscovered popularity, especially in the sophisticated Quinean form of 'holistic coherentism'[81].

As the name of the approach immediately betrays, the central idea in coherentism is that epistemic justification can be analysed in terms of the notion of 'coherence'. The idea is that justification is a function of some relationship between beliefs, none of which are of privileged epistemic starus (as in foundationalism). This 'relationship' is then understood in terms of consistency, probabilistic support and explanatory power. That is, the web of beliefs must avoid internal inconsistencies, form a probabilistically inter-supporting structure and be able to provide good explanations of phenomena that call for an explanation[82].

Unlike the epistemic elitism of foundationalism that bestows certain beliefs with privileged epistemic status ('basic' perceptual beliefs, a priori first principles etc.) and thereby runs into 'the discernibility problem', coherentism is epistemically egalitarian. It does not bestow privileged epistemic status to any sort of beliefs, but takes all beliefs to have equal epistemic standing. All beliefs are justified as long as they avoid inconsistency and are inter-supported by other beliefs and, equally, all beliefs are fallible and hence susceptible to revision. As Quine (1953:43) has notoriously put it '…no statement is immune to revision'.

If a belief (or its inferential implications) does not contradict other beliefs (or their inferential implications) of the overall 'web of belief' and it is also inter-supported by other beliefs, then the belief is unquestionably justified and can be endorsed by

[81] For another defence of a version of holistic coherentism see D.Davidson (2008). Note that Sellars, although a severe critic of foundationalism was not a holistic coherentist, as it is sometimes thought. For he (1997: 79) famously says: 'One seems forced to choose between the picture of an elephant which rests on a tortoise (What supports the tortoise?) and the picture of a great Hegelian a serpent of knowledge with its tail in its mouth (Where does it begin?). Neither will do'.
[82] See J.Kvanvig (2007).

the agent who adheres to coherentism. With 'the inter-supporting web of belief' architecture of justification, coherentism is often (in philosophical consciousness) figuratively compared with the structure of a raft where the planks are all at the same level inter-supporting each other, tied with durable threads from the one edge to the other[83].

Coherentist positions can be broadly distinguished into linear and holistic. The linear coherentist claims that there is a linear chain of individual beliefs that justify each one of our justified beliefs. But linear coherentism has largely been out of favour and philosophers sympathetic to coherentism don't usually attempt to spell out an account of coherentism along these tracks.

It has been rather unpopular because it seems to lead to an uncomfortable dilemma: either to an unpalatable infinite regress of individual beliefs or to a narrow and, therefore, profoundly vicious circularity. This is the case because we either have to push back the chain of individual justifying beliefs ad infinitum or we have to break this infinite regress with a narrow justifying loop. That is, we have to justify one of the supporting individual beliefs with one of the beliefs that were initially in need for justification and, thus, come full circle.

But as with infinitism, an infinite regress of justifications is psychologically impossible for cognitively finite beings like us and, therefore, unpalatable and a narrow justifying loop seems entirely question-begging. For it is question-begging to justify a belief p by reference to a belief q and then justify belief q with the belief p again[84]. For these reasons, I follow the standard practice and hasten to sideline linear coherentism with the intent to turn my attention to the more interesting and popular holistic coherentism[85]. Linear coherentism is currently out of favour and, thus, I unashamedly choose to save my time and space and don't even take the pains to test it again 'the Moorean/Humean lesson' (but this can be left as a conceptual exercise to the fastidious reader).

Let us now turn to the more interesting and popular holistic coherentism. The holistic coherentist claims that what justifies each one of our justified beliefs is not

---

[83] See E.Sosa (2008).

[84] That was Hume's (1986) response to the pragmatic justification of induction by reference to its inductive success. Arguing that inductive beliefs are justified because they tend to get things right and then justify that they get things right because they are inductive is entirely question-begging.

[85] Compare J.Kvanvig (2007: section 1): 'Since the primary examples of coherentism in the history of the view are holistic in nature, I will focus in the remainder of this entry on this version of the view'.

just a chain of other individual beliefs that either leads to an infinite regress or to a narrow circle, but *the whole 'web of our beliefs' at once*. A justified belief stands or falls with the whole 'web of belief' that inherits its justified status. Thus, holistic coherentism, arguably, does not meet either of the problems of linear coherentism. It is neither committed to infinite regress nor to a narrow vicious circularity.

But it is unavoidably committed to wide circularity, though, holistic coherentists claim that there is a world of difference between narrow and wide circularity. Narrow circularity is supposed to be vicious while wide circularity is supposed to be virtuous. Narrow circularity is vicious because the justification loop appeals to only one individual belief at a time and this cannot hold the burden of justification on its own because, as we have seen, it is entirely question-begging.

In contrast, wide circularity is considered to be virtuous because the justification loop appeals to the whole 'web of belief' at once and this, allegedly, can hold the burden of justification on its own; or at least this is what holistic coherentists think. The issue of circularity has been a persistent point of criticism against coherentism, but I won't press it here because I think the Moorean/Humean lesson can adequately do the job and put coherentism to rest[86].

The question now is how holistic coherentism fares against 'the Moorean/Humean lesson'. Is a belief (holistically) coherent justified? Again, the result is I think disappointing. First, comes 'the semantic halt step'. It seems again that there is an inherent semantic gap of 'open feel' that prevents the identification of justification with (holistic) coherence. From the inherent semantic gap of 'open feel' an agent can take the second 'defeating step' and proceed to exercise his imagination and adduce counterexamples.

For instance, think of the biased historian, anthropologist or sociologist that has internalized a Marxist world-view, has a broadly Marxist 'web of beliefs' and is always keen on devising Marxist explanations of historical events and social phenomena. He believes, let us say, that crime is the outcome of the rise of the money-mongering capitalist economic system that sacrifices everything, even human lives, to the altar of its thirst for money. His belief is surely consistent and inter-suported by the rest of his Marxist 'web of belief' and helps providing a neat

---

[86] For a defence of moral coherentism against the circularity objection D.Brink (1989).

explanation of a phenomenon asking for an explanation.

However, this does not really make it justified. Intuitively, crime cannot be solely explained by appeal to just economic reasons and the emergence of capitalism. Surely, there are other psychological, sociological, historical, biological etc. reasons that play a role and this is indicated from the fact that crime has appeared both in societies in pre-capitalist eras and in socialist socioeconomic systems in the capitalist era. Thus, holistic coherence is not sufficient for justification.

The same conclusion is obtained with Hume's 'is/ought argument'. No matter whether a belief is holistically justified, the agent may legitimately take advantage of the inherent semantic 'open feel' and resist endorsement of that certain belief. For instance, a Marxist agent might acknowledge that the belief that 'the proletariat revolution is imminent' is holistically coherent with the rest of his web of beliefs but still resist endorsement of the belief, perhaps, because he suspects that the belief lacks evidential support. Thus, holistic coherentism also fails to raise high enough to meet 'the immunity requirement'.

What the inherent semantic 'open feel' -through the Marxist world-view counterexamples- exploits is what it is often called *'the isolation problem'*[87]. Like the exploitation of 'the discernibility problem' for foundationalism, the exploitation of 'the isolation problem' allows counterexamples to pop up and shoot down holistic coherence. As the example implies, the isolation problem basically points out that no matter how consistent is our 'web of belief', we might be 'cut off' from reality and fail to reliably 'get things right' in the world. In other words, coherence does not seem to be sufficient for justified beliefs[88]. Reaching ideal Rawlsian 'reflective equilibrium' does not on its own put much weight on having justified beliefs that reliably 'get things right' in the world[89].

The reason seems to be that coherence is a purely *internal* affair that neglects how the external world might be. Different coherent theories can be devised for explaining certain phenomena, although these theories, ironically, due to coherence

---

[87] The problem is pressed, among others, by J.Pollock (1985), J. McDowell (1994) and J.Kvanvig (2007).

[88] Again, holistic coherentism could fail in the other direction too, as there could be justified beliefs that are not coherent with one's web of beliefs. This would show that coherence is not even necessary for justification. But in any event, insufficiency alone shows that coherence considerations do not reduce epistemic justification.

[89] I allude here to J.Rawls' (1999) version of moral coherentism.

considerations cannot all be true. As one might put it, different theories are 'underdetermined by the same phenomena'. That is, different theoretical explanations are called in to account for the very same phenomena[90].

A Freudian, for example, might explain the existence of criminal behaviour by appeal to, let us say, unresolved conflicts between the psychological mechanisms of 'the ego, the superego and the id'. A Nietzschean might appeal to the 'will to power', the instinct of dominating others and so on. Yet, the three stories are inconsistent and they can't all be true (though, perhaps, each might contain grains of truth that could be synthesized in a more complex theoretical explanation).

With the discussion of holistic coherentism to an end, I conclude that coherentism, in the light of 'the Moorean/Humean lesson', fails to satisfy 'the immunity requirement' and, thus, does not successfully analyse epistemic justification. The next theory we consider is a virtue theoretic approach. Virtue theories exactly appeared on the theoretical map as an attempt to overcome the unsuccessful efforts of the more traditional approaches of foundationalism and coherentism[91]. The approach we consider is an influential, broadly internalist approach, inspired by Aristotle's virtue ethics: L.Zagzebski's (1996) virtue epistemology.

According to Zagzebski (1996:241):

'A justified belief is what a person who is motivated by intellectual virtue and who has the understanding of his cognitive situation a virtuous person would have, might believe in like circumstances. An unjustified belief is what a person who is motivated by intellectual virtue, and who has the understanding of his cognitive situation a virtuous person would have, would not believe in like circumstances'. Such intellectual virtues include 'intellectual carefulness, perseverance, humility, vigour, flexibility, courage, and thoroughness, as well as open-mindedness, fair-mindedness, insightfulness, and the virtues opposed to wishful thinking, obtuseness, and conformity. One of the most important virtues, I believe, is intellectual integrity' (1996:155).

The notion of virtue is then glossed along broadly Aristotelian lines:

'A virtue, then, can be defined as a deep and enduring acquired excellence

---

[90] See E.Sosa (2008).
[91] See L.Zagzebski (1996) and E.Sosa (2008).

of a person, involving a characteristic motivation to produce a certain desired end and reliable success in bringing about that end. What I mean by motivation is a disposition to have a motive; a motive is an action-guiding emotion with a certain end, either internal or external' (1996:137).

Let us measure Zagzebski's account against 'the Moorean/Humean lesson'. Is justified belief 'what a person who is motivated by intellectual virtue and who has the understanding of his cognitive situation a virtuous person would have, might believe in like circumstances'? Once again, this analysis is off the mark. The first thing to strike the eye is 'the semantic halt step'. It is to notice the inherent semantic 'open feel' that halts any quick identification of epistemic justification with 'what a person motivated by intellectual virtue and with understanding of his cognitive situation a virtuous person would have, might believe in like circumstances'. The second 'defeating step' is to proceed to exploit this inherent semantic 'open feel' and adduce counterexamples. For instance, a belief might be motivated by open-mindedness and vigour but, nonetheless, might not be justified.

Think, for instance, of the medieval alchemist who is open-minded because he takes seriously the possibility that can discover 'the elixir of life' (or transmute dust to gold), he is intellectually careful because takes all steps in an attentive and thorough way and is courageous, patient and perseverant, because he keeps trying despite constant failure. But in spite of the virtuous motivation and his virtuous understanding of his cognitive situation, his belief in the discoverability of 'the elixir of life' is unjustified. It is unjustified because wishful thinking and not evidence drives the conviction that alchemist practices can discover 'the elixir of life'. It seems then that motivation by intellectual virtue in conjunction with the understanding of the cognitive situation a virtuous agent would have is insufficient for epistemic justification.

What the inherent semantic 'open feel' -through the medieval alchemist counterexample- exploits is what we may call *'the virtue insufficiency problem'*. The idea is that we can keep adding epistemic virtues, but this long conjunction of epistemic virtues will still be insufficient for an analysis of justification. No matter how many or which intellectual virtues we add, there will be a semantic gap that will allow counterexamples to emerge. With all the epistemic virtues of the world

involved, an epistemic agent can still not be able to form justified beliefs.

The same conclusion is obtained with Hume's 'is/ought argument'. No matter whether a belief is 'what a person who is motivated by intellectual virtue and who has the understanding of his cognitive situation a virtuous agent would have, might believe in like circumstances', the agent might legitimately resist endorsement of that certain belief. Take, for instance, the alchemist again. The alchemist might be motivated by intellectual virtue to endorse the belief that 'the elixir of life' can be discovered and yet resist endorsement of this belief because he suspects that his whole alchemist errand is driven by wishful thinking, self-deception etc. rather than evidence.

In the face of counterexamples, we conclude again that Zagzebski's virtue theory does not satisfy 'the immunity requirement'. Like (classical and non-classical) foundationalism and holistic coherentism, it does not succeed to deliver a successful semantic analysis of the property of epistemic justification.

Interestingly, however, the failure of the internalist approaches is, I think, *instructive* because all these approaches stumble on problems flowing from the very same source. The source of the 'discernibility, isolation and virtue insufficiency problems', respectively, for foundationalism, holistic coherentism and Zagzebski's virtue epistemology is the abject failure to stably connect the epistemic agent's beliefs with the external world and reliably 'get things right'. They seem to fail to disengage epistemic justification from the internal mental states and character traits of an agent (what beliefs entertains and/or his epistemic virtues) and engage the external world and this seems to spoil their attempts to reliably 'get things right' *in* the world.

Quite independently of what one finds self-presenting, rationally intuitive or explanatorily coherent, or what epistemic virtues motivate the endorsement of a certain belief, the belief in question might fail to be justified and the world might be very different from what these certain beliefs purport to depict. But epistemic justification seems to be truth-conducive and, therefore, an adequate analysis of justification should not just be a function of the internal states and character traits of an agent, simply, because these internal states might be 'cut off' from the external world and thus be very often misguided.

This cues the end of our criticism of internalist approaches. We have found that they are plagued by 'the Moorean/Humean lesson' and fail to satisfy 'the immunity requirement' for a successful semantic analysis of the property of epistemic justification. Interestingly, we also found their failure instructive because these internalist theories seem to exactly fail due to problems flowing from the same spring. Namely, they seem to fail to somehow establish a reliable and stable link between the epistemic agent's beliefs with the external world. We can now turn to externalist theories.

3.3 The Moorean/Humean Lesson Applied to Externalist Theories

The most prominent exponent of externalism has, perhaps, been A.Goldman (1986, 1991, 2002). A.Goldman has done the most work to propound and refine an externalist theory of justification and his work has been to some degree influential. The presentation of his externalist theory will mostly follow his by now classic paper 'What is Justified Belief?'(1991), though, he has been emending and polishing his theory ever since[92]. But the kernel idea of his externalist approach escapes unmolested from his constant emendations and, hence, we need not follow all his constant refinements.

The kernel idea of Goldman is that epistemic justification might be reducible not to a property of beliefs (self-presentation, holistic coherence, motivation by epistemic virtues) one has 'ready at hand' access to (as internalists are assuming), but to reliable belief-forming cognitive processes that function reliably, quite *independently* of whether one has 'ready at hand access to' it or not. Such reliable belief-forming processes are supposed to be perception, memory, inferential reasoning, introspection, even a priori justification (as he argued later in (2002)).

Goldman explicitly sets out for a reductive analysis of epistemic justification, for he does explain that (1991:106) 'I want a set of substantive conditions that specify when a belief is justified… [in] necessary and sufficient conditions… I want a theory of justified belief to specify in non-epistemic terms when a belief is justified…'. Goldman starts from the intuition that unjustified beliefs must share a property in virtue of which are unjustified and, equally, justified beliefs must share a property in virtue of which they are justified. This is basically what in section 2.1 we called 'the

---

[92] See for example A.Goldman (1986).

realist ontological assertion': 'There must be an epistemic justification property in virtue of which justified beliefs are justified'. Then he goes to identify what these justification-conferring properties are.

But let us allow Goldman to explain his train of thought on his own (1992:113):

'…confused reasoning, wishful thinking, reliance on emotional attachment, mere hunch or guesswork, and hasty generalization. What do these faulty processes have in common? They share the feature of *unreliability*: they tend to produce *error* a large proportion of the time. By contrast, which species of belief-forming (or belief-sustaining) processes are intuitively justification-conferring? They include standard perceptual processes, remembering, good reasoning and introspection. What these processes seem to have in common is *reliability*: the beliefs they produce are generally true. My positive proposal, then, is this. The justificational status of a belief is a function of the reliability of the process or processes that cause it, where (as a first approximation) reliability consists in the tendency of a process to produce beliefs that are true rather than false' (Goldman's own emhasis).

He sharpens this kernel idea in various ways and he finally arrives at the following (1992: 124):

'Person *S* is *ex ante* justified in believing *p* at *t* if and only if there is a reliable belief-forming operation available to *S* which is such that if *S* applied that operation to his total cognitive state at *t*, *S* would believe *p* at *t*-plus-delta (for a suitably small delta) and that belief would be *ex post* justified'.

Goldman understands the technical distinction between ex ante/ex post justification in this following way. *Ex ante* justification is used in cases where the agent does not yet endorse a belief, (perhaps, because he is suspending judgement) but reflects on whether a certain belief is justified or not. *Ex post* justification is used in cases where the agent does endorse a belief, yet still reflect on whether the belief is truly justified or unjustified.

With this brief outline of Goldman's process reliabilism, let us now measure his analysis against 'the Moorean/Humean lesson'. Goldman's jargon of technicality aside, is a belief induced by a reliable belief-forming process justified? The first

thing to observe is 'the semantic halt' that impedes any thoughtful and clear-headed agent from a quick reduction of epistemic justification to reliable belief-forming processes. Then the semantic gap stirs 'the defeating step' and all the agent has to do is to exercise his imagination for counterexamples.

The old story repeats itself and counterexamples are looming just round the corner. For instance, think of the miserly man that, though, he has a generally very reliable cognitive process of memory (e.g. remembers where he puts his stuff, easily memorizes poems, bank PIN numbers etc.), perhaps even the formidable memory of a mnemonist, when it comes to remembering his financial obligations, due to his emotional attachment to the accumulation of money etc., is prone to mistakenly remembering that has fulfilled them[93]. This is a case where the generally reliable process of memory seems to go astray and produces a belief we would count as unjustified. Analogous counterexamples where a generally reliable cognitive process goes astray can be devised for perception, introspection and inferential reasoning.

The same result is delivered by Hume's 'is/ought argument'. Any belief produced by generally reliable belief-forming cognitive processes can be resisted. Take an example from Sellars' (1997:37-9) critique of sense-data theories of perception and 'the myth of the (perceptual) Given': John the sales' consultant of a necktie shop. John 'has learned the use of colour concepts in the usual way, with one exception… he has never looked at an object in other than standard conditions'. But then electric lighting is invented. 'His friends and neighbours rapidly adopt this new means of illumination, and wrestle with the problems it presents. John, however, is the last to succumb. Just after [electricity] has been installed in his shop, one of his neighbours, Jim, comes in to buy a necktie.'

'Here is a handsome green one' says John.

'But it *isn't* green' says Jim, and takes John outside.

'Well,' says John, 'it was green in there, but now it is blue.'

'No,' says Jim, 'you know that neckties don't change their colour merely as a result of being taken from place to place.'

'But perhaps electricity changes their colour and they change back again in

---

[93] Such examples occur in real life. S.Freud has described how some of his patients behaved in this way. If I am not mistaken, a case like the one I use as an example has been described in his *The Psychopathology of Everyday Life*.

daylight?'

'That would be queer kind of change, wouldn't it?' says Jim.

'I suppose so' says bewildered John. 'But we *saw* that it was green *in there*.'

'No, we didn't see that it was green in there, because it wasn't green in there, and you can't see what isn't so!'

'Well, this is a pretty pickle,' says John. '*I just don't know what to say*.'(Sellars' emphasis).

The moral of Sellars' story is that no matter how things appear to be to an agent with a generally reliable perception, he cannot deduce that he ought to endorse the belief that they *are* the way they appear to be[94]. I conclude, therefore, that Goldman's process reliabilism fails because it runs into counterexamples and thereby does not satisfy 'the immunity requirement'.

Let us now turn our attention to E.Sosa's (1991, 2007) externalist virtue reliabilism. Sosa's externalist virtue theory is not like Zagzebski's (1996) internalist virtue theory that bases justification on what a person motivated by intellectual virtue and who has the understanding of his cognitive situation a virtuous person would have, might believe in like circumstances[95]. Instead, Goldman-style it bases justification on the reliability of belief-forming 'virtues' where virtues are understood in terms of cognitive competence or ability. Virtues are considered to be relatively reliable and deep-seated intellectual excellences of character.

His initial introduction of intellectual virtue runs as follows (1991:138-139):

'Let us define an intellectual virtue or faculty as a competence in virtue of which one would mostly attain the truth and avoid error in a certain field of propositions F, when in certain conditions C… a faculty or virtue would normally be a fairly stable disposition on the part of a subject relative to an environment'.

His later work still moves in the same trajectory. As he says:

'In holding a certain belief you are foundationally justified in the virtuous

---

[94] Rather, for Sellars (1997: 39): '…this is the heart of the matter. For to say that a certain experience is a *seeing that* something is the case, is to do more than describe the experience. It is to characterize it as, so to speak, an assertion or claim, and – which is the point I wish to stress- to *endorse* that claim' (Sellars' emphasis).

[95] An important difference is that Zagzebski, following Aristotle, takes the intellectual virtues to be largely acquired through moral education. In contrast, Sosa does not seem to pay much attention to the learning aspect.

way to the extent that you are then justified because in so believing you manifest a certain epistemic competence…' (2007:51) and

'The attraction or belief is justified because it is competent… [t]his is why those rational mechanisms are intellectual *competences* because they systematically lead us aright. All seemings delivered by such competences are thereby epistemically justified' (2007:59-60) (Sosa's own emphasis).

For the last time, let us measure Sosa's virtue reliabilism against 'the Moorean/Humean lesson'. Is a belief formed by a reliable belief-forming competence (or virtue) justified? The answer is in the negative for the very same reasons Goldman's broadly similar reliabilism failed. First, there is 'the semantic halt step' that undermines any quick identification and then 'the defeating step' where we can exercise our creativity and imagination and adduce counterexamples.

Reliable belief-forming competences like perception, memory, introspection, reasoning etc. often produce beliefs that cannot be seen as epistemically justified by any means. Think, for example, of the fan with a reliable visual mechanism whom is nonetheless prone to unconsciously overlook referee's mistakes in a football game insofar these are in favour of his team and against the rival team. He is also unconsciously prone to mistakenly attribute wrong decisions to the referee insofar as these are in favour of the rival team and against his own team.

This might happen because the agent's desire for seeing his team winning the game has overwhelmed his desire for truth and perhaps, also, because perception is theory-laden. That is, what you see is being to some extent influenced by the background theory and beliefs, desires, emotions etc. you have. At any rate, what matters here is that although such beliefs are being produced by a reliable visual mechanism they surely do not count as justified. Therefore, such counterexamples show that Sosa's proposal cannot meet 'the immunity requirement' for a reductive analysis.

Hume's 'is/ought argument' delivers the same conclusion. For example, even if a belief is produced by a reliable belief-forming cognitive competence, an agent can resist the normative injunction that he ought to endorse that belief. Think, for example, of cases of inferential irrationality that undeniably beset all reasoning agents (others more, others less). Wishful thinking is a paradigmatic case of such

inferential irrationality. Often, agents with a reliable belief-forming competence of reasoning are seduced by wishful thinking either to draw unjustified conclusions or to resist justified conclusions. Often, on the one hand, we irrationally infer what we desire to infer and not what we ought to infer or, on the other hand, we irrationally resist inferring what we ought to infer because it does not comport with what we desire to infer.

These incidents might occur because we are somehow disposed to unconsciously protect our feelings, self-esteem, sense of dignity etc. We unconsciously protect ourselves from being traumatized by the acknowledgement of either a legitimate but painful inferential outcome of certain of our beliefs or illicitly infer what we ought not to infer from certain of our beliefs just in order to assuage a worry, protect our self-conceited egoism etc.

Now, an agent might resist endorsement of a belief produced by his generally reliable reasoning competence, if he suspects that the belief is based on wishful thinking, self-deception etc. Take the example of the unhappy mother that her soldier son has been declared 'missing in action' during a war concluded ten years ago. The distraught mother infers that her son is probably still alive, held hostage by the enemy, and yet she finds the courage and wisdom to realize that what inclines her to infer that her son is still alive are her motherly instincts of love and affection. In this case, although a reliable belief-forming process produces the belief that her son is probably still alive, the mother resists endorsement of the belief because she realizes that it is motivated by wishful thinking.

In the light of this discussion, I conclude that Sosa's externalist virtue reliabilism -like Goldman's process reliabilism- does not seem to make much progress. More generally, the corollary of our examination of externalist analyses of epistemic justification (Goldman's process reliabilism, Sosa's virtue reliabilism) is that they fall short of a satisfactory reduction by semantic analysis of epistemic justification. They fail to satisfy 'the immunity requirement'. That is, they fail to provide a successful analysis immune to counterexamples and, as a result, they fail to provide a successful semantic analysis of the property of epistemic justification (in necessary and sufficient conditions).

Interestingly, however, like the failure of internalist approaches, the failure of

externalism is *instructive*. While internalists seem to neglect and marginalize the external dimension in justification-conferring, the externalists neglect and marginalize the internal and assume that, for the most part, our belief-forming processes are reliable and do 'mirror' the external reality. The internalist rejoinder here is that this externalist assumption is just an unsupported article of good faith. This is an assumption for which cogent reasons must be given, if it is to be taken seriously, and not just be granted 'on the cheap'.

The overall moral of the application of 'the Moorean/Humean lesson' on internalist and externalist reductionist approaches is that both sides of the divide seem to lack something essential for a successful analysis of justification. Intuitively, an analysis of epistemic justification must, on the one hand, supply some relatively stable and reliable link with the external world but, on the other hand, this relatively stable link must somehow- at least often- be placed *within* the internal. That is, the internal mental states and character traits of an agent (what beliefs entertains and/or his epistemic virtues).

How this is to be done remains unclear, perhaps even enigmatic. But, for the habits-endorsement expressivist account that we pledged to present in Ch.7, there is nothing unclear or even enigmatic because this task is *illusory*. It is simply a fool's errand that cannot be fulfilled. There cannot be a successful analysis of epistemic justification in necessary and sufficient conditions that, as a result, bridges reliably the internal with the external and confers epistemic justification. That is, what we believe (or what is 'in the head') with what is likely to be in the world. The search for such an analysis is only a futile 'wild goose chase'.

The 'wild goose chase' for such an analysis is initially stirred – and this is *the crux of the matter*- from the adoption of 'the referential semantics assumption': justification assertions and attributions purport to refer to a certain property of epistemic justification. From there the analytic naturalistic reductionist moves on to also embrace 'the existential assertion', 'the realist ontological assertion' and, finally, 'the naturalistic analysability assertion'.

As we shall see, for the expressivist all the trouble can be avoided by renouncing 'the referential semantics assumption' that motivates this too stringent and, hence, illusory project in the first place. But again we make a long jump forward because

this is not yet the right time for this discussion.

Besides, just because analytic naturalistic reductionism does not fare well, it does not mean that the entire rich and resourceful theoretical tradition of referential semantics is misguided. And in any case, although there are some subtle and staunch defenders of traditional semantic analysis as the proper philosophical method, not all philosophers in our days believe that semantic analysis is the appropriate philosophical method[96]. Surely, there are other more plausible ways to spell out a referentialist approach that does not cast doubt on *the cornerstone* of their project: the referential semantics assumption.

These remarks signal the end of the argument against naturalistic reductionism. The reached conclusion is that, in the light of 'the Moorean/Humean lesson', probably, no analytic naturalistic reductionist approach will ever succeed. Let us now turn to the conclusion and summary of the chapter.

3.4 Conclusion and Summary of the Argument

In this chapter, I have argued that analytic reductionist approaches fail to reduce epistemic justification. In sections 3.2-3.3, I stepped forward to apply 'the Moorean/Humean lesson' with devastating results for some of the most influential and popular (internalist and externalist) reductionist accounts of epistemic justification. Foundationalism, coherentism, process reliabilism and Zagzebski's and Sosa's virtue approaches have suffered from this application. We, thus, concluded that analytic naturalistic reductionism is, perhaps, not a very promising theoretical project and therefore 'the naturalism-denial assertion' (P3), namely, that 'justification assertions and attributions do not purport to refer to a natural property of epistemic justification' maybe should be seen as plausible.

But as some philosophers sympathetic to referential naturalism will hasten to object, this conclusion is way too quick. To play the devil's advocate, this is what they would probably complain about: 'You make a high jump from the implausibility of analytic naturalistic reductionism to the implausibility of referential naturalism. But this is not quite right. The irreducibility of epistemic justification to a natural property does not impugn reference to a natural property of epistemic justification.

---

[96] Notable defenders of semantic analysis as the proper philosophical method include M.Smith (1994), F.Jackson (1998) and G.Bealer (1998), though, they are often under attack e.g. by S.Lawrence and E.Margolis (2006).

Irreducibility would be fatal for referential naturalism only if analytic naturalistic reductionism was the only option for referential naturalism, but actually is not. Analytic naturalistic reductionism is committed to some obsolete semantic assumptions that can be jettisoned without a nascent theory abandoning its broadly referentialist and naturalistic character'.

In the next chapter, we examine exactly such a theory claiming that we can have referential naturalism without analytic reductionism. This more sophisticated naturalistic approach exploits 'the synthetic naturalism turn' of the 60's and 70's and the pioneering work of S.Kripke and H.Putnam in semantics, epistemology and ontology in order to build a theory that dodges the problems of analytic naturalistic reductionism, though, remains to the core referentially naturalistic.

On how this dialectical manoeuvre can be done and whether it can save the day for referential naturalism, you should turn to the next page. 'Naturalistic reference without reduction by semantic analysis' will be the motto under which synthetic naturalistic reductionism will have a last stand for the cause of referential naturalism. But as a last stand, despite the heroic defence, referential naturalism will eventually succumb; or so I will argue.

Chapter 4 The Problems of Synthetic Naturalistic Reductionism

4.1 Introduction

In this chapter we continue our critical exploration of referential naturalism. But this time we raid a different species of referential naturalism: synthetic naturalistic reductionism (H.Putnam 1981; D.Brink 1989; N.Sturgeon 2007; R.Boyd 2007; H.Kornblith 2005, 2007). This sort of referential naturalism is equipped with more sophisticated dialectical armour than old analytic naturalistic reductionism. Sophisticated dialectical armour exactly prepared to shield referential naturalism from the semantic problems of analytic naturalistic reductionism that, as we saw in Chs.2-3, are being exposed by 'the Moorean/Humean lesson'. This is *the key driving idea* of the project of synthetic naturalistic reductionism. It is to deploy sophisticated ideas that defuse 'the Moorean/Humean lesson' and rescue referential naturalism from ignominy.

Synthetic naturalistic reductionists (often called 'Cornell realists') advocate a form of referential naturalism. Their *core insight* is that the normative property in question exists in nature and we do can refer to it but, crucially, we can't reduce it by a priori semantic analysis. We can't reduce it by a priori semantic analysis because it could only be reduced in terms of an a posteriori identity by means of scientific inquiry[97]. This core insight, as I will explain in the next section 4.2, is meant to defuse 'the Moorean/Humean lesson' and rescue referential naturalism from ignominy.

The theoretical approach of synthetic reductionusm has grown out of, on the one hand, the semantic problems of analytic naturalistic reductionism (brought to the open in Chs.2-3 with 'the Moorean/Humean lesson') and, on the other hand, what we may coin 'the synthetic naturalistic turn' in philosophy that largely took place in the

---

[97] Also, some synthetic reductionists do no just leave open the logical possibility of an a posteriori reduction of a normative property to a natural kind, but go on to firmly postulate the existence of such a natural kind property. H.Kornblith (2005, 2007), for example, asserts that knowledge is a natural kind. Nothing turns on this difference and in the following discussion I will ignore it.

70's. 'The synthetic naturalistic turn' emerged principally out of the groundbreaking work of S.Kripke (1981) and H.Putnam (1975, 1981) in philosophy of language (reference, semantics etc.). Kripke and Putnam criticized traditional 'descriptive theories of reference' as inadequate to accommodate our intuitions about reference to natural kinds and proper names and, instead, introduced 'causal-historical theories of reference'.

The introduction of the causal-historical theories of reference was quite influential (though, as usual, not uncontroversial) and brought a wave of important (some would say revolutionary) changes in our way of thinking about semantics and language and, through their ramified implications, to epistemology, metaphysics and mind[98]. This wave of important changes opened some new and interesting logical space where subtle dialectical manoeuvres could take place with the intent to save the day for referential naturalism[99].

This is what synthetic reductionists shrewdly grasped. They grasped the opening of this new logical space and envisaged that we could exploit causal-historical theories of reference in order to defuse the problematic semantic commitments of analytic naturalistic reductionism and construct a new theory, a theory unburdened by these problematic semantic commitments. Indeed, in the face of the semantic problems of analytic naturalistic reductionism exposed by 'the Moorean/Humean lesson', they set out to exploit the newly opened logical space by causal-historical theories of reference and launch a revamped form of referential naturalism. Yet, despite the sophisticated emendations, this new theory would still remain firmly within the camp of referential naturalism.

Synthetic reductionists' novel and sophisticated ideas were first introduced and applied exclusively on moral properties (H.Putnam 1981; D.Brink 1989; N.Sturgeon 2007; R.Boyd 2007). Although at first sight there is nothing inhibiting an application of these ideas to epistemic (or other) properties, these philosophers weren't really concerned with such an application. But other philosophers where, indeed, concerned with the application of these novel ideas to epistemic discourse. These other

---

[98] For some of the standard objections to causal-historical theories of reference see W.Lycan (2006:Ch.6).
[99] Compare G.Sayer-McCord (1997a: 280-1): 'Naturalists in ethics have found hope in these developments. If property identity does not require synonymy, then whatever the force of the open question argument may be, it does not establish that our moral terms fail to refer to natural properties'.

philosophers in a similar vein have taken up analogous ideas and applied them to epistemic discourse (H.Kornblith (2005) on knowledge; and in some way J.Pollock (1985) on epistemic justification;[100]).

In what follows, I will speak as if the synthetic reductionist ideas have been specifically applied to the property of epistemic justification, although this -as far as I know- has not yet been attempted. Kornblith (2005, 2007) has applied synthetic reductionism to knowledge (rather than epistemic justification) because he takes 'the knowledge first view' and Pollock's (1985) account of epistemic justification is not strictly speaking a synthetic reductionist approach (see n.100 on why).

This much about how synthetic reductionism came to the fore, its core insight and the key driving idea behind this core insight. The current chapter presents and examines the argument of synthetic reductionism. This is the plan for this presentation and examination: in section 4.2, I present synthetic naturalistic reductionism in some detail. I introduce its theoretical commitments and explain how these theoretical commitments purport to defuse 'the Moorean/Humean lesson' and rehabilitate a sophisticated form of referential naturalism. Afterwards, in section 4.3, I critically examine the plausibility of the theory.

In section 4.3, I argue that despite the employment of sophisticated semantic assumptions by synthetic reductionists with the intent to defuse 'the Moorean/Humean lesson', an equally sophisticated version of Moore's classical 'open question argument' can be constructed to show that their efforts do not make much progress. This equally sophisticated version of 'the open question argument' is 'the epistemic twin earth argument'[101].

Afterwards, in sections 4.4-4.5 I introduce, examine and reject a subtle kind of normative realism developed by G.Sayre-McCord (1997a) which concedes that 'the epistemic twin earth argument' shows that normative properties are not natural kinds,

---

[100] Pollock's (1985) approach is not strictly speaking a synthetic reductionist approach. While he argues that epistemic justification is irreducible by semantic analysis (1985:123), he does not appeal to causal-historical theories of reference and a posteriori property identities in order to launch a viable form of referential naturalism about epistemic justification. But he does seem to hint to that general direction when, in passing, he talks about epistemic norms as natural kinds and the conceptual roles of epistemic concepts discovered by 'a posteriori conceptual analysis'(1985:168-175).

[101] A sophisticated Humean 'is-ought argument' can be constructed in the same 'twin earth' pattern the Moorean 'open question argument' follows. For reasons of simplicity, though, in section 4.3 I forgo this theoretical construction because the sophisticated Moorean 'open question argument' is already complex enough and, in addition, can tackle the issue on its own. But this theoretical construction can be left as a conceptual exercise to a willing reader.

but still aspires to resist the antirealist conclusion Timmons and Horgan (2007) draw from it. Sayre-McCord's argument is concerned with moral properties but I conveniently reinterpret the argument in terms of epistemic properties, as nothing hinges on this reinterpretation. He calls his theory 'normative kinds realism' and, accordingly, I call the epistemic version of the theory 'epistemic kinds realism'. Finally, in section 4.6, I conclude this chapter with a summary of the argument so far and prepare the ground for our next Ch.5.

4.2 Synthetic Naturalistic Reductionism Introduced

In this section, we present the theoretical commitments of synthetic naturalist reductionism and explain how they are meant to rescue referential naturalism from the semantic problems exposed by 'the Moorean/Humean lesson'. We make explicit the theoretical commitments of the approach and explicate how these theoretical commitments are supposed to perform the gambit to defuse the semantic problems of analytic naturalistic reductionism, as exposed by 'the Moorean/Humean lesson', and rehabilitate referential naturalism. This is our *twofold task* in this section.

Let us now see the theoretical commitments of synthetic naturalistic reductionism and how these are meant to defuse 'the Moorean/Humean lesson'. Like analytic naturalistic reductionism, synthetic naturalistic reductionism is a referentialist approach. It starts from the implicit adoption of 'the referential semantics assumption' (P2), namely, accepts that justification assertions and attributions purport to refer to a certain property of epistemic justification. Unlike fictionalists, though, he also accepts that not only our justification assertions and attributions purport to refer to such an epistemic justification property but that this property exists.

The synthetic reductionist thinks that an epistemic justification property exists because assumes that at least some of our allegedly justified assertions and attributions are, indeed, justified. This is the familiar from section 2.2, *'the existential assertion'*. But if there are at least some justified beliefs, he continues, then there must be a property of epistemic justification in virtue of which these justified beliefs are justified. This is the also familiar from section 2.2, *'the realist ontological assertion'*.

At this juncture, however, synthetic and analytic naturalistic reductionism part

roads. The synthetic reductionist thinks that this property of epistemic justification that must somehow be 'out there', if at least some of our beliefs are to be justified, could *not* in principle be reduced by semantic analysis to a certain natural property. Unlike analytic naturalistic reductionism, the synthetic approach vehemently denies that in principle we can reduce by semantic analysis the property of epistemic justification. This is his particular stance because, in the light of 'the Moorean/Humean lesson', he accepts the anti-reductionist verdict of Ch.3 and takes the property of epistemic justification to be irreducible by semantic analysis.

In short, he scraps -what in Ch.2 we dubbed- 'the naturalistic analysability assertion': 'The epistemic justification property (in virtue of which justified beliefs are justified) can, in principle, be reduced by semantic analysis to a certain natural property'. Instead, he assumes that the property of epistemic justification is, in principle, irreducible by the traditional means of semantic analysis. Let us call the assumption that the property of epistemic justification is, in principle, irreducible by semantic analysis to a natural property, *'the naturalistic unanalysability assertion'*.

The synthetic reductionist discards 'the naturalistic analysability assertion' for 'the naturalistic unanalysability assertion' because, *crucially*, unlike the analytic naturalistic reductionist, he does not adhere to 'a descriptive conception of reference'. In contrast, the analytic naturalistic reductionist with 'the naturalistic analysability assertion' does adhere to a descriptive conception of reference. That is, he assumes that there is a directly descriptive natural property 'out there' reducing epistemic justification and this property can be discovered by semantic analysis[102].

In turn, the descriptive conception of reference assumes -what D.Brink (1989: 162) has called- 'the semantic test of property identity'. Namely, that we can discover property identities (and, hence, accomplish reduction) only by means of traditional semantic analysis. In other words, according to 'the semantic test of property identity', reduction is a matter that lies in the hands of property identities discovered by semantic analysis.

The synthetic reductionist, however, does not agree with any of these three

---

[102] The use of 'directly descriptive' needs some explication. With 'a directly descriptive natural property' I mean a property that can be perceived simply by means of ordinary macroscopic eyesight e.g. being a tree. In this sense, properties perceived through microscopic devices are 'indirectly descriptive'. They are descriptive of the underlying microstructure of a certain natural property e.g. water is H2O.

assumptions: 'the naturalistic analysability assertion', 'the descriptive conception of reference' and 'the semantic test of property identity'. In the footsteps of Kripke and Putnam, he takes a dim view of these assumptions. On the one hand, he renounces 'the naturalistic analysability assertion' for 'the naturalistic unanalysability assertion' and, on the other hand, abandons the descriptive conception of reference for a causal-historical conception of reference. That is, he doesn't think that directly descriptive features of the world regulate the use of normative sentences.

Rather, he thinks that features of the world that are not in any sense directly descriptive causally regulate the use of such referential sentences. Subsequently, he exploits the causal-historical conception of reference in order to show that 'the semantic test for property identity' is obsolete and then, based on this point, defuse 'the Moorean/Humean lesson' and rescue referential naturalism. Let me elaborate on their train of thought.

First, very briefly, let us sketch how the two different conceptions of reference are meant to work. According to the traditional descriptivist conception of reference, we can refer to natural kinds (and proper names), let us say, instances of gold by identifying at least some of its main directly descriptive characteristics like its yellowish colour, soft texture, shininess, density, malleability and ductility[103]. But for the causal-historical conception of reference, identifying directly descriptive characteristics of a natural kind is utterly inadequate to regulate reference to the natural kind in question. It is inadequate because these directly descriptive features do not *uniquely identify* instances of gold. In other words, they don't reduce the property of being gold.

Arguing by counterexample, causal-historical theorists press that there are cases where a different natural kind bears these directly descriptive features but is not actually gold. As one might know from some good western film, 'fool's gold' (or iron pyrite) also has these directly descriptive characteristics but is not actually gold. These counterexamples seem to show that no matter which directly descriptive characteristics of a natural kind we might identify these might prove deceptive and, hence, badly regulate reference to that certain natural kind. We might use 'gold' to

---

[103] There are different descriptive theories of reference in the theoretical marketplace (see G.Frege (1997), B.Russell (1956), J.Searle (1958)) but these differences are inconsequential to our discussion here and, therefore, are conveniently overstepped.

refer to things (e.g. iron pyrite) associated with these directly descriptive characteristics but, nevertheless, these things do not really be instances of 'gold'. As Shakespeare has said, 'all that glitters is not gold'[104].

In contrast, for the causal-historical theorist, what -causally- regulates referring sentences to gold is the property of being the element on the periodic table with atomic number 79. The rationale for this proposal is straightforward. The property of being the element with the atomic number 79 is what *uniquely identifies* and, hence, reduces instances of gold. In the light of this unique identification, we can say that sentences referring to the property of being 'gold' (e.g. My ear rings are made of gold) are actually referring to the property of 'being the element with atomic number 79'. The use of these referential sentences is causally regulated by the underlying essential property of gold.

Now, the adoption of the causal-historical conception of reference is what is meant to defuse 'the Moorean/Humean lesson' by rendering obsolete *the hidden semantic assumption* that lies at the heart of 'the Moorean/Humean lesson': 'the semantic test of property identity'. 'The semantic test of property identity' assumes that only through a priori semantic analysis we can reduce a property to another. We can reduce a property to another if and only if we discover a certain identifying property by semantic analysis. But as Kripke and Putnam taught us, this is a way too restrictive view of reduction because we can have a posteriori property identities as well. Property identities discovered by the empirical means of scientific inquiry (and not by traditional 'armchair' semantic analysis) that causally regulate the reference of the corresponding terms.

Take, as another example, the natural kind of silver. According to the Kripke-Putnam picture, the property of 'silver' is reducible to the property of 'being the element on the periodic table with atomic number 47'. This is an a posteriori property identity discovered through scientific inquiry and does not depend in any way to any analysis of the meaning of the notion of silver. No matter how hard we try, this is not something discoverable by 'armchair philosophizing'. It is not, because the property of being the element with atomic number 47 is not in any sense inherent in the meaning of the concept of silver and, hence, we can't derive it by

---

[104] This appears somewhere in his *The Merchant of Venice.*

meticulous semantic analysis. This is, after all, why we have a property identity that does not rest on synonymy unveiled by a priori analysis of meanings.

In a parallel way, the causal-historical conception of reference is then used to defuse the semantic problems of analytic naturalistic reductionism (as exposed by 'the Moorean/Humean lesson'). In the light of a causal-historical conception of reference, synthetic reductionists claim that as the property of, let us say, being 'H20' reduces and causally regulates the property of being 'water', there could be a certain natural kind property reducing and causally regulating the property of epistemic justification (and other normative properties). That is, there could be a natural kind property of epistemic justification causally regulating the use of justification assertions and attributions. Call this claim *'the causal-historical reference assertion'*.

For it might be the case that normative properties, in spite of the inherent 'open feel' semantic intuitions sabotaging reductive semantic analyses, do have a posteriori property identities that have not yet been discovered. The parallel with natural kinds is sharply clear. For, surely, before Lavoisier's discovery in the 1750s that water is H2O, if we were asked whether water is H20, an inherent semantic 'open feel' would indeed accompany the reduction. But there was after all an a posteriori property identity showing that the question is actually definitively closed and this explained away the 'open feel' semantic intuitions. The same could be the case with the property of epistemic justification (and other normative properties).

The inherent semantic 'open fell' intuitions could accompany reductive efforts by semantic analysis because there is no such directly descriptive natural property that can accomplish such a feat. These semantic intuitions could constitute an indication that epistemic justification is not reducible by a priori semantic analysis. But as in the case of water, these semantic intuitions will disappear if we discover an a posteriori property identity reducing epistemic justification. Thus, this proposal, evidently, seems to explain away the inherent semantic 'open feel' intuitions that accompany attempts to reduce the property of epistemic justification to a natural property (or to deduce a prescriptive 'ought to believe' from a referential 'is').

This concise discussion yields a general overview of the synthetic reductionists' theoretical commitments and how these are meant to defuse the semantic problems

of analytic naturalistic reductionism exposed by 'the Moorean/Humean lesson'. But before moving on to rehearse in more succinct and explicit terms the synthetic reductionist's theoretical commitments, we should pause to make clear a last point.

It is important to be stressed that synthetic reductionists only claim that we can't reduce the property of epistemic justification by the traditional means of a priori semantic analysis. These philosophers, though, do not suggest that the property of epistemic justification somehow *must* be reducible by the sophisticated means of a posteriori scientific inquiry[105]. Besides, their proposal rests on the mere logical possibility of the property of epistemic justification being reduced by means of scientific inquiry, not to any necessary reduction (in the by now familiar pattern of a posteriori property identities).

In any case, the synthetic naturalistic reductionist proposal does not hinge on any necessary reduction by a posteriori means of the epistemic justification property to a certain natural property. It is meant to work even on the assumption that we will *never* reduce by empirical means the property of epistemic justification. For, their core insight is that the property of epistemic justification *could* be a natural kind and we do can refer to it. But if such an a posteriori reduction will be somehow accomplished in the future, then synthetic reductionists will welcome this because it will *vindicate* their theoretical project. It will show that they were right to propose that the property of epistemic justification could be reduced to a natural kind in an a posteriori pattern.

This short digression for clarifying this last point finished, in more succinct and explicit terms these are the commitments of synthetic naturalistic reductionism:

• The Referential Semantics Assumption: 'Justification assertions and attributions purport to refer to a certain property of epistemic justification'

• The Existential Assertion: 'There are at least some justified beliefs'

• The Realist Ontological Assertion: 'There must be an epistemic justification property in virtue of which justified beliefs are justified'

• The Naturalistic Unanalysability Assertion: 'The epistemic justification

---

[105] Compare: 'The moral realist *may* choose to agree that goodness is *probably* a physical property but deny that it has any physical definition, analytic or other' R.Boyd (2007:173; my emphasis) and 'Naturalistic identity claims should be construed on the model of other common identity claims, such as water=H2O... [m]oral properties can *be* natural properties, though, even if they are not identical with natural properties' D.Brink (1989: 157; my own emphasis).

property (in virtue of which justified beliefs are justified) cannot, in principle, be reduced by semantic analysis to a certain directly descriptive natural property'

- The Causal-Historical Reference Assertion: 'Justification assertions and attributions, in principle, purport to refer to a certain natural kind property (whatever this may be) causally regulating their use '

We have now completed our twofold task for this section. That is, on the one hand, we presented the theoretical commitments of synthetic naturalistic reductionism and, on the other hand, explained how they are meant to defuse the semantic problems of analytic naturalistic reductionism, as exposed by 'the Moorean/Humean lesson', and rehabilitate a sophisticated form of referential naturalism. Let us now turn our attention to the critical examination of these proposals. The other part of critical examination follows in section 4.4.

4.3 Synthetic Naturalistic Reductionism and Epistemic Twin Earth

In section 4.2, we introduced the theoretical commitments of synthetic naturalistic reductionism and explained how they exploit the causal-historical theory of reference with the intent to defuse the semantic problems (exposed 'by the Moorean/Humean lesson') of analytic naturalistic reductionism and articulate a viable form of sophisticated referential naturalism.

In this section, we present our critical examination of the proposal. We argue that the deployment of sophisticated dialectical weaponry by synthetic naturalistic reductionism with the intent to defuse 'the Moorean/Humean lesson' does not seem to make much progress. It does not seem to make much progress because it runs into an equally sophisticated version of 'the open question argument': 'the epistemic twin earth argument'. As we shall explain, it appears to succumb to 'the epistemic twin earth argument'.

The 'epistemic twin earth argument' is inspired from the analogous 'moral twin earth argument' of M.Timmons and T.Horgan (2007). The 'moral twin earth argument' has been propounded in a number of papers and has been specifically used against so-called synthetic naturalistic reductionism's application to moral properties[106]. Here, since our discussion concerns the property of epistemic justification, I reinterpret the argument in epistemic terms. But we should bear in

---

[106] See M.Timmons and T.Horgan (2007) and M.Timmons (1999: ch.3).

mind that there is no substantial dialectical difference between the initial moral and the current epistemic interpretation. I broadly follow Timmons' and Horgan's (2007) original argument.

M.Timmons and T.Horgan (2007), following H.Putnam's (1975: 215-271) renowned 'twin earth thought experiment' for causal-historical theories of reference (and semantic externalism), constructed an analogous thought experiment purported to launch a strengthened, non-classical 'open question argument' that could confront synthetic reductionism head-on. The idea was to build a modernized 'open question argument' that founders synthetic reductionism in the same way Moore's classical 'open question argument' founders analytic reductionism.

In concise terms, this is 'the epistemic twin earth argument'. Like in Timmons and Horgan's moral version of the argument, consider two planets: 'good old earth' and 'epistemic twin earth'. The two planets are alike in all respects: geography, surroundings etc. and, significantly, its inhabitants, respectively, 'earthlings' and 'twin earthlings' both speak English. They have the same conceptual repertoire (i.e. vocabulary) and use the epistemic predicates 'is justified', 'is rational', 'is well-grounded' etc.

The *only* difference is that in 'good old earth' epistemic predicates are causally regulated, as a matter of some a posteriori property identity, by some coherentist property (e.g. Quinean holistic coherentism) while in 'epistemic twin earth', as a matter of some a posteriori property identity, epistemic predicates are causally regulated by some foundationalist property (e.g. Chisholmian self-presentation). The question now is how we should understand this sole difference between 'good old earth' and 'epistemic twin earth'.

As Timmons and Horgan (2007) suggest, there are two hermeneutic options:

- We could say that 'earthlings' and 'twin earthlings' *differ in meaning and reference,* since epistemic predicates are causally regulated by different natural properties in each case: a coherentist for 'earthlings' and a foundationalist for 'twin earthlings'.

- We could say that, despite the causal regulation by different natural properties, their disagreement is *not one of meaning and reference* but rather *one of genuine disagreement* in epistemological theory and belief.

The first hermeneutic option seems to be the case in Putnam's (1975: 223-227) original 'Twin earth argument'. In Putnam's argument, the term 'water' in planet earth is causally regulated by H20, while in planet twin earth 'twin-water' is causally regulated by a property, alas, unbeknown to us earthlings, the XYZ. In such a case, Putnam (1975) argued, 'earthlings' and 'twin earthlings' seem to mean and refer differently with the term 'water' and 'twin-water': H20 in planet earth and XYZ in planet twin earth.

Very briefly (and rather unfairly to the richness and complexity of Putnam's discussion), this is the case because, as Putnam cogently argued, 'the twin earth thought experiment' shows that psychological states are not sufficient on their own to determine meaning, though, they are probably necessary for meaning. They are not sufficient on their own to determine meaning because, as the thought experiment shows, intuitively, people with the same psychological states can have different meanings. Earthlings with 'water' mean H20 and Twin earthlings with 'twin-water' mean XYZ.

Earthlings and Twin earthlings have different meanings because the external world, importantly, *does* contribute to the meaning of our terms and expressions. That is, meaning is at least partially determined by the external world. It is determined by social conventions constantly informed by the dissemination of scientific advances and discoveries. Therefore, no one can reasonably deny that at least partially, as philosophers of mind say, 'mental content is broad'. In Putnam's (1975:227) famous colloquial phrase, 'Cut the pie any anyway you like, meanings ain't in the head!'; or, at least, not *entirely* in the head (as that would mean sliding to a form of Berkeleyan idealism!).

In contrast, like Timmons and Horgan (2007) in their 'moral twin earth argument', I opt out for the second hermeneutic option. And this is *the heart of the matter*. I contend that, unlike Putnam's original 'twin earth thought experiment', in the 'epistemic twin earth thought experiment' we seem to have rather different hermeneutic intuitions. Intuitively, we don't seem to think that there is difference in meaning and reference as in Putnam's original scenario, but rather genuine disagreement in epistemological theory and belief.

Genuine disagreement in epistemological theory and belief that appears irresolvable, even in the logically possible scenario that we did have an a posteriori

property identity that reduces the property of epistemic justification. For, it appears that even if we had a posteriori property identities for epistemic properties, intuitively, we would *still* genuinely disagree about these properties. In contrast with natural kind terms, where a posteriori property identities seem to resolve genuine disagreement and definitively close the question, in the case of epistemic (and other normative) properties potential discovery of a posteriori property identities does not seem to promise the same closure results. It does not seem to promise resolution of genuine disagreement and definitive closure of the question.

So, while the question: 'Is the natural kind of water reducible to H2O?' in the light of an a posteriori property identity intuitively appears closed, the analogous question: 'Is epistemic justification reducible to a certain natural kind property?', even in the light of a presumed a posteriori identity, would *still* intuitively appear indeterminately open. Like Timmons' and Horgan's (2007) original argument, therefore, I conclude that sophisticated synthetic reductionism does not seem to make much progress. Unfortunately, despite its subtlety and sophistication, it still founders on revived inherent semantic 'open feel' intuitions. It still runs into another devastating 'open question argument'. Call this lesson '*the epistemic twin earth lesson*'.

However, as in the case of Moore's classic 'open question argument' we should be careful about the exact logical import of 'the epistemic twin earth argument'. It should not be mistaken for something that it is not and, surely, it is not a conclusive rebuttal of synthetic naturalistic reductionism. Very briefly, both arguments, Moore's classical 'open question argument' and 'the epistemic twin earth argument' seem to fall prey to W.Frankena's (1939) famous 'begging-the–question objection'[107]. As have seen in section 2.3, Moore's argument presupposes that we can't reach analytic naturalistic definitions and, since this is what is really at stake, it is enough to 'beg the question'.

For, despite our unsuccessful attempts at least since Plato, we can't inductively preclude that we won't find one in the future. Timmons and Horgan's argument analogously presupposes that even if we reach a posteriori identities causally regulating epistemic terms the semantic 'open feel' intuitions will remain and this,

---

[107] See A.Miller (2005: 167-8) for some discussion.

again, *pre*judges what is really at stake. For, after all, the semantic 'open feel' might vanish in thin air when we finally discover the a posteriori property identities of epistemic terms. Given then that Moore-style 'open question arguments' are committed to such inductive prejudgements they are doomed to beg the question against referential naturalism (analytic or synthetic) and, hence, they are far from being conclusive.

But, as in the case of Moore's classical 'open question argument', many people are inclined to think that the argument strikes something deep that should not be missed simply in virtue of its inconclusiveness. Besides, no one would deny that many inconclusive arguments do create a strong presumption for their case. The argument seems to strike something deep because it captures the entrenched intuition that, no matter what happens, genuine disagreement about epistemological theory and belief is irresolvable. Therefore, many philosophers are willing to accept the argument as enthymatic. That is, as *an inference to the best explanation* for the existence and persistence of these inherent 'open feel' semantic intuitions. The inherent 'open feel' semantic intuitions persistently arise because there is no a posteriori property identity reducing epistemic justification.

With 'the epistemic twin earth lesson in hand', in the next section 4.4, we introduce Sayre-McCord's (1997) 'epistemic kinds realism' that aspires to accept 'the epistemic twin earth lesson' but still resist its antirealist conclusion. Again, it does not fare very well.

4.4 Epistemic Kinds Realism Introduced

G.Sayre-McCord's (1997a) interesting argument for 'normative kinds realism' comes as a response to Timmons' and Horgans' (2007) 'moral twin earth argument'. As we have seen, 'the moral twin earth argument' amounts to a sophisticated 'open question argument' devastating synthetic naturalistic reductionism. Yet, Sayre-McCord's intention is to show that there is still open logical space for a form of moral realism. It is to show that while Timmons' and Horgans' (2007) argument against synthetic naturalistic reductionism correctly traces something important, they nevertheless hasten to draw the wrong conclusion.

That is, by the fact that we seem to have 'open feel' semantic intuitions about a possible reduction of moral properties to natural kinds, we should not hasten to lead

to the hands of moral antirealism/expressivism (as Timmons and Horgan urge). We can still be moral realists and accept 'the epistemic twin earth lesson', namely, that normative properties are not reducible to natural kinds. Now, as in the case of synthetic reductionists, Sayre-McCord's argument is in moral terms and, therefore, henceforth I freely reinterpret it in epistemic terms.

Sayre-McCord (1997a) argues that we can accept 'the epistemic twin earth lesson', namely, that in all evidence there are no natural kinds reducing epistemic properties and still resist Timmons' and Horgan's (2007) conclusion that there are no epistemic properties. We can achieve this much, Sayre-McCord (1997) contends, by drawing a subtle and delicate distinction between 'natural kinds terms' and 'epistemic kinds terms'. Crucially, epistemic kinds terms are supposed to operate 'much like' natural kinds terms but not exactly like them (and this is supposed to open the new logical space). Let me explain the difference between natural and epistemic kinds terms and how this is supposed to open new logical space that absorbs 'the epistemic twin earth lesson' and still rescues epistemological realism.

As we have seen in section 4.2, according to 'causal-historical theories of reference' natural kinds terms (like water, gold, silver etc.), in the light of our best scientific theory, are causally regulated by certain natural properties e.g. Water is H20. These natural properties a posteriori reduce and, therefore, uniquely identify the respective natural kinds. Since we have reduction and unique identification of the natural kinds by certain natural properties, we can legitimately talk of causal regulation of the use of the natural kind terms by the respective natural properties. That is, we can talk of a certain causal-historical link between the correct use of the natural kind terms and the natural property reducing the term in question.

Now, importantly, Sayre-McCord argues that the case with epistemic kinds terms (justification, rationality etc.) is just slightly different from natural kinds terms; though, this slight difference makes an (epistemic) world of difference. The difference is that, unlike natural kind terms, epistemic kind terms are *not* supposed to be causally regulated by certain natural kind properties. This is what Timmons and Horgan (2007), Sayre-McCord (1997a) concedes, correctly traced with their 'epistemic twin earth argument'.

For, if epistemic kinds terms were to be causally regulated by a certain natural

kind property -according to our best scientific theory- as synthetic reductionists envisaged, then in the 'epistemic twin earth argument' we wouldn't have 'open feel' semantic intuitions. We wouldn't feel that no matter which natural property a posteriori reduces an epistemic kind term, there would still be irresolvable normative disagreement about normative theory and belief. But in fact, as we have canvassed in section 4.3, we do have such 'open feel' semantic intuitions.

Unlike natural kinds terms, epistemic kinds are meant to be causally regulated only by *instances* of the kind, not by the kind itself. We then proceed to observe and track the similarities of the instances of the relevant epistemic kinds and form a theory of what they could mean. Afterwards, our resultant moral theory 'fixes the reference' of the epistemic kinds in the same way our scientific theory 'fixes the reference' of natural kinds (with Kripkean 'rigid designators'). Once the reference of epistemic kinds is fixed, this gradually becomes part of linguistic conventions and epistemic kinds terms come to mean what the fixation of reference takes them to mean.

With the distinction between 'natural/epistemic kinds terms' at hand, Sayre-McCord then argues that the discovery of the true nature of epistemic kinds is an unfolding 'process' that takes time, reflection and responsiveness to new information about such epistemic kinds. This new information, though, could come only from the maturation of normative epistemological theory, not scientific inquiry. It could come from the process of development, maturation and fruition of an adequate epistemological theory through due argument, reflection and new information. Scientific advances do not have a say in this process, at least not in any direct and important way.

Worthy of special notice is also that Sayre-McCord's (1997a: 271) proposal is supposed to remain ontologically neutral. That is, it remains neutral on what these epistemic kind properties are from an ontological point of view (natural or nonnatural)). As far as his proposal is concerned, they might turn out to be either natural or nonnatural and his proposal remains open to both interpretations. This is a crucial assumption and we will back to this in our criticism of Sayre-McCord's argument.

But here is a nice quote that captures much of the essence of Sayre-McCord's

(1997a:269-270) proposal in his own words:

> 'Discovering the referents of our [epistemic] terms is a process, I'll suggest,
> of discovering what normatively significant kinds –and not causal-
> explanatorily significant kinds- regulate our [epistemic] beliefs. The
> regulation involved must be, I'll nonetheless assume, a causal regulation,
> so the kinds in question will end up being causally relevant at least to the
> explanation of our use of the terms. Yet when an [epistemic] kind is
> appropriately implicated in the explanation of our use of some term it will
> not be because we causally interact with the kind itself. Rather, it will be
> because (i) our use is causally responsive to what are, in fact, instances of
> the kind and (ii) the use to which the term is put is one of referring to
> whatever normatively significant kind it is that they are instances of… Of
> course just what might count as appropriate causal regulation is itself a
> complicated and controversial matter. I assume that a crucial part of the
> story will highlight features of the use of the term in question that make its
> use responsive to new information about the kind in question'.

According to Sayre-McCord, once the distinction between the 'natural kinds semantics' of synthetic reductionism and the 'epistemic kinds semantics' is crisply drawn, we can harvest important theoretical fruits. On the one hand, new logical space is cleared for a form of epistemological realism that defuses 'the epistemic twin earth lesson' and, on the other hand, certain important explanatory advantages are secured. Let me explain how epistemic kinds semantics accomplishes this much.

First, 'epistemic kinds semantics' accept 'the epistemic twin earth lesson' because these semantics are not causally regulated by a natural kind (according to our best scientific picture of the world). Rather, they are causally regulated by instances of epistemic kinds that, subsequently, our epistemological theory exploits in order to fix the reference of epistemic kinds terms. The 'epistemic twin earth argument' neglected this available logical space because identified causal regulation with causal regulation by a natural kind and then hastened to pronounce their antirealist verdict: there are no epistemic properties.

But as Sayre-McCord points out, we need not take this step. We can settle for causal regulation by instances of the kind. Thus, presumably, for Sayre-McCord

'open feel' semantic intuitions exist because our best epistemological theory has not yet discovered the nature of epistemic kinds. Therefore, 'the epistemic twin earth lesson' applies to the 'natural kinds semantics' of synthetic reductionism but not to 'epistemic kinds semantics'.

Second, 'the [epistemic] kinds approach also helps to make sense of why no robust analytical definitions are available' (1997a: 269). That is, it makes sense of why 'the Moorean/Humean lesson' worked so efficiently. The '[epistemic] kinds approach' makes sense of why 'the Moorean/Humean lesson' worked so efficiently 'because it treats proposed 'definitions' of epistemic terms as discoveries not settled by current use, prevailing linguistic conventions, or declarative stipulation…' (1997a:269). The true nature of an epistemic kind is quite different of what we currently think it is and discovering this nature is 'a process'; not something discovered with the easiness and speed of the blink of an eye.

Third, with the 'epistemic kinds semantics' we can still vindicate our pre-theoretical realist intuitions. We can make sense of talking, on the one hand, about the same epistemic properties when we use epistemic terms and, on the other hand, about facts of the matter that can resolve epistemic disputes. Unlike 'the epistemic twin earth argument', we can opt for the first hermeneutic option and interpret normative disagreement as disagreement about meaning and reference (as in Putnam's original scenario), not as irresolvable disagreement about normative theory and belief.

This vindicates the realist intuition that when we disagree about epistemic matters we disagree about *something*; we are not just talking past each other (this is the familiar 'realist ontological assertion'). There is, presumably, a fact of the matter that can settle the dispute. These 'facts of the matter' we purport to refer to when we use epistemic terms are determined by 'the epistemic kinds' (whatever these may be).

To illustrate this last point, Sayre-McCord (1997a: 285) gives the example of people having different (deontological and consequentialist) conceptions of justice. Let me illustrate the same point by giving an analogous example of people having different conceptions of epistemic justification. Some people might accept a Quinean theory of justification and count as justified only the beliefs that conform to Quine's holistic coherentism, while others may accept a Chisholmian theory of

justification and count as justified only the beliefs that are either self-presenting themselves or based on self-presenting beliefs. Yet, in this case we don't see people as talking past each other, but 'as holding competing views about [justification]'. That is, holding competing views about what justification is. And, surely, when we talk about justification we imply both that the property of justification exists and has a certain determinate nature.

This concludes our presentation of 'epistemic kinds realism'. Let us now turn to the critical appraisal of this attempt to rescue epistemological realism from 'the epistemic twin earth lesson'.

4.5 Epistemic Kinds Realism Criticized

A first quick (yet, too quick) complaint is that Sayre-McCord's argument might break down on the 'much like' analogy with natural kind terms. It might break down on exactly the point of the alleged causal regulation of epistemic kinds terms by instances of the kind, but not by the kind itself. For, we can clearly understand how causal regulation by kinds fixes the meaning and reference of natural kind terms, but is not so clear how causal regulation by instances can do the same for epistemic kinds terms.

'H20' causally regulates the use of 'water' because it reduces and, therefore, uniquely identifies water. But we don't have property identities that reduce, uniquely identify and, hence, can causally regulate the use of epistemic kind terms. That was the problem with synthetic reductionism in the first place. We don't have such property identities at our disposal and 'open feel' semantic intuitions indicate that the prospect of discovering any is unpromising. If this is correct, then epistemic kind terms can only be causally regulated by some sort of natural kind properties (given that analytic naturalistic reductionism is not an option). But this is the initial synthetic reductionist approach and we have seen that it runs into 'an open question argument'. Thus, the argument goes, epistemic kinds semantics do not fair any better than synthetic reductionism's natural kinds semantics.

However, this argument entirely misses the point of Sayre-McCord's theory. For the insight of Sayre-McCord was to stress that we shouldn't -like Timmons and Horgan- assume that the only sort of causal regulation is causal regulation *by a kind*. We can also have causal regulation *by instances of the kind* and this is good enough

for a viable form of epistemological realism.

For although we might not yet know what reduces the epistemic kind of 'justification' and, therefore, can't claim that 'justification' is causally regulated by a certain natural kind property, we can still speak of 'justification' as being causally regulated by instances of this kind property; or, more accurately, at least what we currently think are instances of this kind property. We can still be 'causally responsive' to instances of what we take to be, let us say, justified beliefs, theories, explanations etc. even if we don't know which property reduces 'justification'.

We can then systematize the similarities of what we take to be instances of 'justification' and come to form a speculative theory of what 'justification' could be. This process could take place time and time again as our speculative theories might collapse in the face of counterexamples, recalcitrant intuitions etc.

But there is no reason to be overly pessimistic about the possibility that one day this constant formulation and trial of more and more refined theories won't arrive at the true nature of what justification is. And once epistemological theory settles on what 'justification' really is (or any other epistemic kind), we can then 'fix the reference' of our epistemic term and, gradually, this will become part of the established linguistic convention. Yet, it takes epistemological theorizing and not scientific inquiry for this task and epistemological theorising needs time and effort.

Think, for example, how we gradually came to realize epistemic facts like the truth-conducivity of inductive experimentation and meticulous observation of natural phenomena. Nowadays, this sounds a commonplace truism to our modern ears but actually it was the result of a slowly unfolding process that gave rise to the scientific method sometime around 16th-17th century. In pre-scientific communities, the formulation from scratch of animistic explanations (and other 'just so stories') of natural phenomena was (and in some places still is, I am afraid) a common epistemic practice. The realization of such epistemic facts was exactly a socio-historical process, not something that suddenly dawned on us. That is, it was a process that took argument, reflection, controversy etc.

In conclusion, this preliminary objection does not seem to make much progress against epistemic kinds realism as it exactly misses its key insight, namely, the distinction between causal regulation by kinds and causal regulation by instances of

kinds.

A second more promising objection to Sayre-McCord's theory is the following. Given the ontological neutrality of Sayre-McCord's theory, it faces a dilemma: either epistemic kinds are natural properties or nonnatural properties. Let us take the first horn of the dilemma. If epistemic kinds are natural properties, then we should be able to discover them either by semantic analysis or by a posteriori scientific methods, as what exists in nature can in principle be discovered.

What exists in nature has independent causal powers and takes part in causal interactions and, sooner or later, we can track and identify such things. Thus, in principle, we should be able to discover the nature of epistemic kinds, even if this discovery is a process that will take time and effort. To suggest otherwise, namely, that there can be natural properties that, in principle, are not discoverable is to mystify nature and I don't think that Sayre-McCord would be happy to embrace this.

But this first horn runs straight into our still powerful 'open feel' semantic intuitions about a potential discovery -by our matured best epistemological theory this time- of the nature of epistemic kinds. We can thus build on these intuitions and level another 'open question argument' in the form of an 'epistemic twin earth argument'.

Once again we have 'earth' and 'twin earth'. As usual, the two planets are identical in all imaginable respects (surroundings, geography, language etc.). There is only one difference. The use of 'earthlings'' notion of 'justification' is causally regulated by the coherentist property discovered by Quine. Quine has won the hearts of earthlings and they came to realize that some form of holistic coherentism captures the true nature of justification. They proceeded to 'fix the reference 'of justification according to Quine's coherentist dictates and by now with justification they mean what Quine taught them to mean.

But in contrast, the use of the notion of justification in twin earth is causally regulated by a property discovered by a twin earthling foundationalist named Chisholm. The twin earthling foundationalist has again won the hearts of twin earthlings and they came to realize that some form of 'self-presentation' captures the true nature of justification. They proceeded to 'fix the reference' of justification and by now with justification they mean what Chisholm taught them to mean.

The question now is how to understand the normative disagreement between earthlings and twin earthlings on 'what justification is'. There are again two hermeneutic options:

- We could say that 'earthlings' and 'twin earthlings' *differ in meaning and reference,* since justification is causally regulated by a different natural property in each case: a Quinean coherentist for 'earthlings' and a Chisholmian foundationalist for 'twin earthlings'.

- We could say that, despite the causal regulation by different natural properties, their disagreement is *not one of meaning and reference* but rather *one of genuine disagreement* in epistemological theory and belief.

Now it seems, again, that we should side with (b). Disagreement about what justification is seems to be genuine disagreement in epistemological theory and belief because, intuitively, we feel that no matter with which property our best epistemological theory identifies justification we will *still* be doubtful about this reduction. And we will feel doubtful because of the inherent 'open feel' semantic intuitions concomitant to such reductive efforts. Thus, the first horn of the dilemma leads to a dead end for epistemic kinds realism [108].

Let us now follow the second horn of the dilemma and take epistemic kinds to be nonnatural properties. If epistemic kinds are nonnatural properties, then epistemic kinds are irreducible to more basic naturalistic terms and our pursuit of analysing their nature through a process of trial-and-error will never discover the nature of epistemic kinds. This nonnaturalist reading is in agreement with what 'the epistemic twin earth lesson' suggests.

But although Sayre-McCord's theory embraces the conclusion of 'the epistemic twin earth lesson', if it is given a nonnaturalist reading, it then meets serious epistemological and ontological problems we will discuss in the next Ch.5. I postpone discussion of these problems until the next chapter but, as we will see in the Ch.5, these problems are important and make nonnaturalist theories seem implausible. Without any third ontological reading option, epistemic kinds realism seems to have been cornered as neither a naturalistic nor a nonnaturalist reading of this theory can really help it.

---

[108] This is, broadly, how Gibbard (2003: 159-168) argues against Sayre-McCord, though, without the conceptual apparatus of a 'twin earth' thought experiment.

Sayre-McCord might still have a bifurcated rejoinder, though. First, against 'the epistemic twin earth argument', he might respond that it works against his proposal *only* if we assume that the discovery of the true nature of epistemic kinds is possible and that such a discovery will entail the evaporation of 'open feel' semantic intuitions and, subsequently, the resolution of entrenched normative disagreement[109]. But this is an assumption that nowhere in the text Sayre-McCord appears to concede and, moreover, at one point seems to explicitly hint that he is not conceding it. For as he says (1997a: 288): 'It is not a forgone conclusion that our [epistemic] terms will find a reflection, let alone a vindication, in the best normative theory'.

This is somewhat brief and cryptic, but he seems to be saying that is not quite clear whether we will ever discover the true nature of epistemic kinds, and even if we do discover it, is also not clear whether this will be vindicated as such. And it might not be vindicated as the true nature of epistemic kinds because the 'open feel' semantic intuitions of agents might still resist its acceptance and, hence, resolution of normative disagreement might not be achieved.

Second, Sayre-McCord might point out that our argument, if correct, seems unable to explain our pre-theoretical realist intuitions -as his theory does- and that this is another defect of taking hermeneutic option (b) above. Perhaps, he might even make the by far stronger claim that any theory that cannot explain these pre-theoretical realist intuitions does not really take off the ground in the first place, because these intuitions are an *a priori constraint* on any possible theoretical explanation of normative discourse.

Sayre-McCord's (imagined) bifurcated rejoinder goes deep. My brief counter-response is this. On the first prong, given that Sayre-McCord does not concede neither the assumption that the true nature of epistemic kinds is discoverable nor that such a discovery will bring evaporation of the 'open feel' semantic intuitions and convergence, 'the epistemic twin earth argument' is, indeed, rendered toothless. It is rendered toothless because he accepts that there are epistemic kinds and we might never discover their nature and, a fortiori, even if we discover their nature we might never come to vindicate it as such.

If we grant to Sayre-McCord this assumption, ultimately, the discussion boils

---

[109] This style of response to 'the open question argument' has been explored by M.Smith (1994).

down to a stalemate of conflicting intuitions. Sayre-McCord has realist intuitions and goes for hermeneutic option (a). That is, he thinks that normative disagreement is one of meaning and reference, even though normative properties are not natural kinds and science is not of much help on normative issues. We disagree about what something is; we don't just talk past each other.

In contrast, I have antirealist intuitions and go for hermeneutic option (b). That is, I think that normative disagreement is not one of meaning and reference but rather one of genuine disagreement in epistemological theory and belief. We genuinely disagree about epistemic kinds because there are no such things with a certain determinate nature; (though, I wouldn't say that disagreeing agents talk past each other just because there are no such epistemic kinds).

However, I don't think that we should be so charitable to grant this assumption to Sayre-McCord. Reasonably, if epistemic kinds are natural properties we should be able to discover them either by semantic analysis or by a posteriori scientific methods, as what exists in nature can in principle be discovered. For, to suggest that something can be natural and, thereby, exist in nature and still, in principle, not be discoverable by scientific inquiry is to mystify nature in an anachronistic manner.

We should be able to discover the nature of epistemic kinds, even if their initial discovery does not lead to immediate convergence. For it is true enough that 'open feel' semantic intuitions obstructing convergence and agreement may persist even when we do discover the truth about some question. Think for example of Copernicus' proposal of heliocentricism. It took some time for people to be convinced that geocentricism is mistaken. Equally, disagreement might persist even when we discover (if ever) the true nature of epistemic kinds.

But, reasonably, we expect rational agents to come bit-by-bit, step by step (through sober reflection, argument, full information etc.) to recognise the truth of the matter at some point as it happened with the historical example of heliocentricism. If we have discovered the truth about a certain question, then we should expect rational agents at some point to acknowledge this truth. And this is mere common sense.

Think, more generally, how science evolves through Kuhnian 'paradigm shifts'[110].

---

[110] I allude to T.Kuhn (1996).

At the beginning there is strong resistance to the revolutionary theory, even vehement outright rejection. But gradually as passions subside and wisdom returns the better scientific theory seems to take the lead. And this happens because we are rather rational beings (though not *all* that rational). We can understand if a theory does better than another in terms of explanatory power, theoretical virtues etc. and finally reason things out. Of course, this might take time, from months to years, decades even centuries in some occasions. Such incidents abound in the history of science.

Here is an example from the history of medicine. The 'Pap test' is an easy, low-cost and reliable method of detection and prevention of cervical cancer and other cytologic diseases of the female reproductive system. Dr G.Papanicolaou discovered it in the early 1920's. Yet although the adduced evidence was clear, it took two decades (and many more female lives) until the incredulous physician's community accept it. The same could happen with a potential discovery of the nature of epistemic kinds. We could discover their true nature and disagreement still persists for some time. But the point is that, if we have discovered what epistemic kinds really are, it is reasonable to expect that the evaporation of 'open feel' semantic intuitions and convergence sooner or later will eventually come.

The second prong goes again deep. It is true that everyday life normative discourse seems to exemplify these realist intuitions and, therefore, a theory should better adequately accommodate these intuitions. I have conceded this much when in section 1.4 said that the great promise of expressivism is to account for the realist pretensions normative language in such an antirealist framework. The problem, however, is that, as we have seen (Chs. 2-4), striving to accommodate these realist intuitions in a realist framework runs us into very serious difficulties. In the light of these difficulties then, it is perhaps (just perhaps) worthwhile to endeavour to accommodate these realist intuitions in an antirealist framework.

Thus, we shouldn't blankly reject even the mere possibility of such an exploration from the very start. We should take it as a working hypothesis that stands in need of conceptual exploration. Whether such an exploration is worthwhile or not, it is something that can only be evaluated in retrospect at the end of the road, after we have gone all the way down in exploration of such a theoretical explanation. It is not

something that should (or can) be decided from the outset.

Besides, Plato himself introduced his nonnaturalist/realist 'theory of forms' in the *Phaedo* as a mere working hypothesis to be judged from the theoretical fruit it can produce, not as the ultimate truth of the matter. If realists can do this then antirealist should be allowed to do the same. And for one thing, we should not forget that philosophy is praised for its open-mindedness and dialectical tolerance.

In conclusion, these considerations entitle us to think that the epistemic kinds approach does not really yield the promised land of a plausible form of epistemological realism. In the next section 4.6, I wrap up the argument of the current chapter and prepare the ground for the next chapter 5.

4.6 Conclusion and Summary of the Argument

In this chapter, we raided two sophisticated species of referential naturalism taking advantage of causal-historical theories of reference in order to defuse the semantic problems of old analytic naturalistic reductionism, as exposed by 'the Moorean/Humean lesson', and rescue referential naturalism. We found that, in spite of the deployment of sophisticated dialectical weaponry and subtle dialectical twists, the synthetic naturalist reductionism and epistemic kinds realism are again subjected to criticism.

In section 4.2, we concisely presented the theoretical commitments of synthetic naturalistic reductionism and how they are meant to defuse 'the Moorean/Humean lesson' and rescue referential naturalism. We explained how by taking advantage of causal-historical theories of reference, synthetic naturalistic reductionists render obsolete the semantic assumption that lies hidden in the very heart of 'the Moorean/Humean lesson', 'the semantic test of property identity', and open logical space for a sophisticated form of referential naturalism.

In section 4.3, we presented our criticism for synthetic naturalistic reductionism. We argued that it runs into a sophisticated 'open question argument', 'the epistemic twin earth', and succumbs immediately to revived 'open feel' semantic intuitions. In sections 4.4-4.5, we introduced and examined a neglected alternative that endeavours to rescue epistemological realism from the teeth of the 'epistemic twin earth lesson': epistemic kinds realism. We found that epistemic kinds realism faces a dilemma: either epistemic kinds are natural properties or nonnatural. If they are natural

properties, they run again into 'the epistemic twin earth lesson'. If they are nonnatural properties, they run into serious epistemological and ontological problems we promised to present in the next chapter 5.

The discussion of synthetic naturalistic reductionism and epistemic kinds realism signals the end of the discussion of referential naturalism and yields its long awaited conclusion. In all evidence, justification assertions and attributions do not purport to refer to a certain natural property of epistemic justification. This is the familiar 'naturalism-denial assertion' (P3) of 'the epistemic justification puzzle' and the discussion of Chs.2-4 seems to buttress its plausibility. Referential naturalism (sophisticated or traditional) does not seem to be particularly plausible.

With this aporetic conclusion at hand, a figure like Socrates would probably drop the exploration of the question 'What is epistemic justification?' here and, perhaps, suggest in his familiar ironic tone that, strangely, although we seem to have justified beliefs we don't know what epistemic justification is[111]. But I hope that Socrates would excuse the temerity and, perhaps, recklessness of my inexperienced youth because I will try to explore some more theoretical approaches to the puzzle. I hope that he would excuse me on the grounds that philosophy is foremost a challenge in conceptual exploration and Socrates valued conceptual exploration more than anything else; or at least this is the picture Plato inherited to us.

With Socrates' eulogies assumed then, in the next chapter our little odyssey of conceptual exploration continues. We examine another venerable theoretical camp that traces its roots to Plato, Socrates' own great disciple: nonnaturalism. That is, the possibility that the property of epistemic justification is some sort of nonnatural, sui generis property. A property that is not reducible to other natural properties and in this sense is primitive, not derivative. The camp of nonnaturalism is also meant to be raided. In the end, 'the nonnaturalism-denial assertion' (P4) of 'the epistemic justification puzzle' will also appear in favourable light.

That is, it will appear plausible that justification assertions and attributions do not purport to refer to a certain nonnatural property of epistemic justification. On how this is supposed to be accomplished, we should turn to the next page. We now set sail

---

[111] As he does, for example, in Plato's (2001) *Lysis* (223b*)*: '…we have made ourselves ridiculous… [f]or these others will go away and tell how we believe we are friends of one another… but what a friend is, we have not yet succeeded in discovering'.

for the lands of nonnaturalism and in the next chapter we present and examine the plausibility of two forms of the nonnaturalist camp: 'naïve' nonnaturalism and J.McDowell's (1994) more sophisticated quietist version of nonnaturalism.

Chapter 5 The Problems of Nonnaturalism

5.1 Introduction

In Chs. 2, 3 & 4 we criticized referential naturalism (analytic naturalistic reductionism, synthetic naturalistic reductionism and -if given a naturalistic reading- epistemic kinds realism). We found that such approaches, on the one hand, run into devastating 'open question arguments' (classical or sophisticated). We concluded that referential naturalism is not very plausible. In terms of 'the epistemic justification puzzle', this is translated as 'the naturalism-denial assertion' (P3) is plausible, namely, justification assertions and attributions do not purport to refer to a certain natural property.

Still, this is not sufficient to undermine the idea that some form of epistemological realism is plausible. It is not sufficient to undermine the idea that there is a property of epistemic justification we purport to refer to when we make justification assertions and attributions. For it might be the case that the property of epistemic justification is a *sui generis* property, a property irreducible to naturalistic properties (analytically or synthetically). Besides, if we do not respect Sayre-McCord's (1997a) professed ontological neutrality about his normative kinds and feel free to ascribe to him the view that normative kinds are nonnatural properties, then we quickly arrive to such a nonnaturalist realist position.

In this chapter 5, we examine this alternative realist tradition that its traces go back to Plato's (middle period) 'theory of forms': nonnaturalism. Nonnaturalists concur that referential naturalism (of any sort) is not very plausible. They embrace our conclusion that 'the naturalism-denial assertion' (P3) is plausible, namely, that justification assertions and attributions do not purport to refer to a certain natural property of epistemic justification. Instead, nonnaturalists propose that justification assertions and attributions purport to refer to a certain sui generis property. A property that is nonnatural and therefore primitive. That is, a property that cannot be reduced (analytically or synthetically) to other more basic natural properties in terms of necessary and sufficient conditions.

A recent notable exponent of such a nonnaturalist approach to epistemic discourse

has been T.Williamson (2000). Williamson (2000), who takes 'the knowledge first view', has argued that knowing is an irreducible, sui generis - in his own terms- 'mental state'. In our terms, this means that knowing is a property irreducible to more basic properties like the traditional true justified belief analysis[112].

As far as I know, though, there is no straightforward example of an epistemologist both taking 'the epistemic justification first view' and claiming that epistemic justification is a sui generis, nonnatural property. In what follows, I assume that there is such an epistemologist in order to examine the plausibility of this open logical space. Of course, the criticism of this open logical space carries over to Williamson's nonnaturalism about knowledge (and other sorts of nonnaturalism e.g. moral, aesthetic, semantic, modal, mathematical etc.).

As I explain in section 5.2, the motivation for a nonnaturalist approach to 'the epistemic justification puzzle' is twofold. On the one hand, a nonnaturalist approach can easily defuse the semantic problems of referential naturalism exposed by 'open question arguments'. It can defuse these semantic problems because it accepts the lessons of the 'open question arguments': 'the Moorean/Humean lesson' and 'the epistemic twin earth lesson'. That is, it concurs that the inexistence of a natural property (of any kind) of epistemic justification best explains the reality of our inherent 'open feel' semantic intuitions.

On the other hand, it can easily explain our pre-theoretical realist intuitions about epistemic discourse. This much is achieved with the postulation of 'the realist ontological assertion': there must be a property of epistemic justification in virtue of which justified beliefs are justified. The postulation of 'the realist ontological assertion' allows us to explain a bunch of pre-theoretical realist intuitions surrounding epistemic discourse. Such realist intuitions are the justification similarity relation, epistemic objectivity and truth, the existence of mind-independent epistemic norms, the possibility of factual epistemic (dis)agreement etc.

But unfortunately, as I explain in section 5.3, although at first sight nonnaturalism seems particularly auspicious, it eventually turns out to be a theoretical giant with

---

[112] Someone might question here why talk of knowing as irreducible to a mental state is equivalent to talk of knowing as irreducible to a property. I think that they are equivalent because if we think that knowing is irreducible to more basic mental states (like true justified belief), then we are committed to think that the property of knowing is also irreducible to more basic properties (like true justified belief).

feet of clay. That is, it seems to have certain problems that seriously undermine the plausibility of the theory. These problems are of an epistemological and ontological ilk. Here is a quick glimpse of what these epistemological and ontological problems are about.

First, there is the 'epistemic access problem', namely, how we could ever track the existence of such a nonnatural property. Second, there is the 'queerness problem', namely, what such a 'queer' sui generis property could ever be. Third, as I explain in section 5.4, nonnaturalism has problems explaining -what I call- the 'epistemic supervenience desideratum', namely, a desideratum asking how a nonnatural property of epistemic justification could ever necessarily supervene on a certain natural property.

These thorny epistemological and ontological problems seem to prompt the conclusion that there is no nonnatural property of epistemic justification. But while such a simple and 'naïve' version of nonnaturalism easily succumbs, J.McDowell (1994, 1998) has defended a subtle neoAristotelian, quietist version of nonnaturalism that aspires to rescue a form of nonnaturalism. In sections 5.5 I introduce and in section 5.6 examine McDowell's 'naturalism of second nature' that aspires to make a more modest version of nonnaturalism naturalistically respectable and find this too wanting, as it fails to stifle the epistemological and ontological worries for nonnaturalism.

Thus, in the end of the examination, I arrive at the conclusion that nonnaturalism does not seem a very plausible answer to the Socratic question 'What is epistemic justification?'. In terms of 'the epistemic justification puzzle', this is translated as 'the nonnaturalism-denial assertion' (P4) is plausible, namely, justification assertions and attributions do not purport to refer to a certain nonnatural property of epistemic justification.

Overall, this is the plan for this sketched critique. First, in section 5.2, I make explicit the theoretical commitments of 'naïve' nonnaturalism and explain how these commitments are supposed to address 'the epistemic justification puzzle' in an elegant fashion. That is, in a manner that does not run into the semantic problems of the referential naturalistic approaches (highlighted in Chs.2-4) and, in addition, explains our pre-theoretical realist intuitions about epistemic discourse.

In sections 5.3-5.4 I criticize the 'naïve' nonnaturalist approach, namely, explain its epistemological and ontological discontents. First, in section 5.3, I explain how 'the epistemic access problem' and 'queerness problem' impugn nonnaturalist approaches. That is, I explain why it is difficult to understand, first, how we could ever come to track such a nonnatural property and, second, how could such a property ever exist.

Second, in section 5.4, I explain how nonnaturalism seems unable to explain 'the epistemic supervenience desideratum'. I initially introduce what is the epistemic supervenience desideratum and then explain why it is difficult to understand how any nonnatural normative property -not just epistemic justification- could ever necessarily supervene on a natural property. In section 5.5, I go on to introduce McDowell's sophisticated 'naturalism of second nature' and in section 5.6 criticize and reject his proposal. Finally, in section 5.7 I preview the argument of the chapter and prepare the ground for the next chapter 6.

5.2 'Naïve' Nonnaturalism Introduced

As it is accustomed, let us first make explicit the theoretical commitments of nonnaturalism and how these are meant to defuse the semantic problems of referential naturalism, capture our pre-theoretical realist intuitions about epistemic discourse and, hence, resolve 'the epistemic justification puzzle' in elegant fashion. Like the camp of referential naturalism, the nonnaturalist camp starts from *'the referential semantics assumption'* (P2). It assumes that epistemic justification assertions and attributions purport to refer to a certain property of epistemic justification.

Like referential naturalists again, nonnaturalists also accept that not only our justification assertions and attributions purport to refer to such an epistemic justification property but that this property exists. The nonnaturalist thinks that an epistemic justification property exists because he assumes that at least some of our allegedly justified assertions and attributions are, indeed, justified. This is *'the existential assertion'*. But if there are at least some justified beliefs, he adds, then there must be a property of epistemic justification in virtue of which these justified beliefs are justified. This is 'the realist ontological assertion'.

Up to this point, referential naturalistic approaches and nonnaturalism go hand

with hand. They are in agreement that some form of epistemological realism should be true. But this 'realism fellowship' does not last because referential naturalists and nonnaturalists dissent about the exact form of this epistemological realism. They disagree about the exact ontological character of the property of epistemic justification.

Nonnaturalism first parts roads with analytic naturalistic reductionism. The nonnaturalist accepts the anti-reductionist corollary of Ch.3, namely, that we cannot reduce epistemic justification by semantic analysis and, therefore, rejects 'the naturalistic analysability assertion'. The nonnaturalist thinks that this property of epistemic justification that must somehow be 'out there', if at least some of our beliefs are to be justified, could not in principle be reduced by semantic analysis to a certain natural property. This is the familiar '*naturalistic unanalysability assertion*'.

The nonnaturalist also parts roads with the synthetic naturalistic reductionist because he rejects 'the causal-historical reference assertion'. He denies that our justification assertions and attributions purport to refer to a certain natural kind property causally regulating their use. He is convinced that, in general, referentially naturalistic approaches cannot reduce epistemic justification (analytically or synthetically) and would therefore gladly embrace 'the epistemic twin earth lesson'. This follows from the more general methodological stance of nonnaturalists.

Nonnaturalists usually adopt a traditional stance towards philosophical methodology and take philosophy to be a clearly distinct and autonomous field of inquiry from science. For nonnaturalists, philosophy is one thing, science is another and empirical inquiry cannot really provide answers to philosophical puzzles. This is what lurks in the background of their scepticism for referential naturalism and boosts their conviction of the naturalistic irreducibility (analytic or synthetic) of normative properties. Call this irreducibility claim '*the naturalistic irreducibility assertion*'. Since referential naturalists can't reduce the property of epistemic justification, they also cannot account for our pre-theoretical realist intuitions and adequately address 'the epistemic justification puzzle'.

Rather, the nonnaturalist thinks that the property of epistemic justification should be taken to be *nonnatural*. He thinks that since there must be a property of epistemic justification in virtue of which justified beliefs are justified and this property cannot

be reduced (analytically or synthetically) to naturalistic terms, then there is no other way out of the puzzle than nonnaturalism. The property of epistemic justification should be considered in principle to be a sui generis, naturalistically irreducible property. This would nicely account for 'the realist ontological assertion', namely, that there must be a property of epistemic justification in virtue of which justified beliefs are justified and, as a result, accommodate a bunch of other pre-theoretical realist intuitions surrounding epistemic discourse. Call this claim *'the nonnaturalist postulate assertion'*.

In more succinct and explicit terms, these are the theoretical commitments of nonnaturalism:

- The Referential Semantics Assumption: 'Justification assertions and attributions purport to refer to a certain property of epistemic justification'

- The Existential Assertion: 'There are at least some justified beliefs'

- The Realist Ontological Assertion: 'There must be an epistemic justification property in virtue of which justified beliefs are justified'

- The Naturalistic Irreducibility Assertion: 'The epistemic justification property cannot, in principle, be reduced (analytically or synthetically) to a certain natural property'

- The Nonnaturalist Postulate Assertion: 'The epistemic justification property must, in principle, be an irreducible (analytically and synthetically) nonnatural property'

Let us now elaborate in more detail how these theoretical commitments are utilised by nonnaturalism. As we already mentioned in section 5.1, the motivation for nonnaturalism is twofold. First, these theoretical commitments defuse the semantic problems of referential naturalism exposed by 'open question arguments'. As we have seen in Chs. 2-4, referential naturalistic approaches run into 'open question arguments' exposing their semantic problems and thwarting their reductive efforts. Analytic naturalistic reductionism runs into a classic Moorean 'open question argument' and synthetic naturalistic reductionism and epistemic kinds realism (if given a naturalistic reading) run into the sophisticated 'epistemic twin earth argument'.

Nonnaturalism easily defuses these 'open question arguments' because endorses

'the naturalistic irreducibility assertion'. It accepts that justification assertions and attributions do not purport to refer to a certain natural property of epistemic justification. That is, it concurs that the inherent 'open feel' semantic intuitions indicate that there is no natural property reducing (analytically or synthetically) epistemic justification and, hence, accepts their anti-reductionist verdict.

Second, these theoretical commitments vindicate our pre-theoretical realist intuitions about epistemic discourse. Such pre-theoretical realist intuitions concern the justification similarity relation, epistemic objectivity and truth, the existence of mind-independent epistemic norms, the possibility of convergence and agreement in the pursuit of an inquiry by rational epistemic agents etc. They vindicate such intuitions because 'the realist ontological assertion' entitles us to do so. Let me explain how this much is achieved.

Pre-theoretical realist intuitions appear to be somehow ingrained in the syntactic (and semantic) structure of our everyday language (and thought). For everyday epistemic discourse seem to be *intentional*: it seems to be about epistemic properties (justification, rationality etc.). That is, justification attributions, rationality assertions like 'S justifiably believes that p', 'p is rational' etc. seem to purport to refer to the respective epistemic properties. This intentionality of epistemic discourse seems to imply that there are mind-independent, objective facts about matters epistemic that if discovered, ideally, could elicit agreement among rational disputants.

These pre-theoretical realist intuitions are actually what drive the emergence of 'the referential semantics assumption', 'the existential assertion' and 'the realist ontological assertion' in the first place. These three theoretical commitments are supposed to be driven in the first place from our pre-theoretical realist intuitions about epistemic discourse. For in everyday life epistemic discourse when we talk about the justification of an individual belief or the justification of a whole bunch of interconnected beliefs called a theory, we seem to be talking about a certain property that this individual belief or theory instantiates. This is what 'the referential semantics assumption' captures.

We also presume that at least some of our beliefs and theories are indeed justified. Almost everyone concedes that we do have at least some justified beliefs, no matter whether these beliefs are ultimately true. This is what 'the existential assertion'

captures. And, presumably, there is a certain property of epistemic justification in virtue of which justified beliefs are justified. There is a common property shared by these beliefs that entitles us to call them by the same name: 'justified'. This is what 'the realist ontological assertion' captures.

The realist ontological assertion then allows us to vindicate a bunch of other pre-theoretical realist intuitions. First, it easily explains the justification similarity relation. For, trivially, if there is a nonnatural property of epistemic justification then all justified beliefs are justified in virtue of the nonnatural property of epistemic justification. Second, it explains the intuitions concerning epistemic objectivity and truth. If there is a nonnatural property of epistemic justification, then it is a matter of epistemic fact whether a belief is justified or not. There is a truth of the matter whether a belief is justified or not. If it instantiates the epistemic justification property then is justified. If not, then it is not justified.

Third, it vindicates the intuition that there is an objective fact about matters epistemic that if discovered, ideally, could elicit convergence and agreement among rational disputants. For given the truism that one ought to believe what is (best) epistemically justified and that we can identify instances of epistemic justification, then ideally we can aspire to convergence between rational disputants.

Finally, it vindicates the intuition concerning the existence of objective, mind-independent epistemic norms. For, it is often considered that epistemic justification is regulated and, hence, it is inter-definable by objective, mind-independent epistemic norms[113]. A belief is justified if and only if is licensed by the correct epistemic norms. That is, the epistemic norms that *are* indeed the correct ones no matter which norms we *think* are the correct ones. But if epistemic justification is inter-definable with the correct epistemic norms and epistemic justification is a nonnatural property, then these epistemic norms should also be considered to be nonnatural.

This much being said about the twofold motivation behind nonnaturalism, it is of critical importance to fully understand why nonnaturalists in various ways do not hesitate to resort to invocation of such a sui generis property and risk the danger of being criticized for unwarranted spookiness and mystery. Nonnaturalists are not, of course, philosophically naïve or something of the sort. Far from it, they are often

---

[113] See J.Pollock (1985:125) and T.Williamson (2000: 242).

very thoughtful and competent philosophers fully aware of this kind of criticism and yet still opt for this kind of theoretical explanation.

Like their forefather Plato (and for some readings Aristotle), contemporary nonnaturalists opt for this kind of theoretical explanation because they are not prepared to give up these theoretical commitments emerging from our pre-theoretical realist intuitions. The reason they are not really prepared to give up these theoretical commitments emerging from our pre-theoretical realist intuitions (or at least question them) is their more general stance on philosophical methodology.

Their methodological take on how to do philosophy is firmly traditional and, thus, conservative. They start from the pre-theoretical realist intuitions ingrained in everyday life thought and talk and aspire -whatever the theoretical price we have to pay- to vindicate these intuitions. That is, to accommodate these intuitions in a certain realist theoretical framework. In Aristotelian terms, they strive to 'save the (normative) phenomena' at any cost.

In the light of this philosophical methodology, they opt for nonnaturalism because they think is the only way, on the one hand, to defuse the semantic problems of referential naturalism exposed by 'open question arguments' and, on the other hand, explain and vindicate our pre-theoretical realist intuitions about epistemic discourse. In short, they think that pre-theoretical realist intuitions must be explained no matter how high is the (epistemological and ontological) price we have to pay and nonnaturalism is the only obvious way to do so.

Arguably, this is what their great forbearer himself, Plato, had in mind when he introduced his nonnaturalist 'theory of forms'. Here is Plato in the *Phaedo* (100c-d), for the first time in the history of Western philosophy introducing (as a working hypothesis) the idea of nonnaturalism:

> 'Now I do not yet understand…nor can I perceive those other ingenious causes. If anyone tells me that what makes a thing beautiful is its lovely colour, or its shape or anything else of the sort, I let all that go, for all those things confuse me, and I hold simply and plainly and perhaps foolishly to this, that nothing else makes it beautiful but the presence or communion (call it which you please) of absolute beauty, however it may have been gained;'.

As I read the text (and Plato in general), Plato here exhibits the usual realist policy of 'vindicating pre-theoretical realist intuitions whatever the price'. He admits that he doesn't understand how the property of beauty could be reduced to a certain natural property of 'lovely colour or shape or anything else of the sort' 'for these things confuse me'. They confuse him because they run into counterexamples, as there are things with lovely colour that are not beautiful or beautiful things without lovely colour etc.[114]. But he holds 'simply and plainly and perhaps foolishly to this' that if we are to vindicate our pre-theoretical realist intuitions about aesthetic discourse, then beautiful things must be beautiful in virtue of the property of beauty itself, 'however it may have been gained'. That is, no matter how this property 'causes' beautiful things to be beautiful, we need to postulate such 'absolute' beauty if we are to make sense of our realist intuitions.

This brief discussion draws an outline of 'naïve' nonnaturalism's commitments and theoretical aspirations. It is concise and somewhat simplifying but I think is sufficient for present purposes. In the next sections we turn to the critical examination of 'naïve' nonnaturalism. As usual, our critical examination of nonnaturalism will have a raiding character. It will be short and direct and will unveil the theoretical giant's feet of clay. The first dose comes in the next section 5.3 and highlights the epistemological and ontological problems of nonnaturalism.

5.3 The Epistemological and Ontological Problems of 'Naïve' Nonnaturalism

Like the theoretical approach of nonnaturalism itself, its epistemological and ontological problems are hardly a contemporary discovery. Nonnaturalism has a long history of facing such thorny problems and these have consistently been considered the primary reason for rejecting nonnaturalist theories (or, at least, viewing them with due perplexity in spite of the other attractions they bear). In Plato's own time, Aristotle (2003a, 2003b, 2006) forcefully pressed these problems and in modern times philosophers from I.Kant (2003), to A.J.Ayer (1946), J.Mackie (1977), D.M.Armstrong (1980), S.Shoemaker (2004) and beyond did the same.

These historically consistent critical reactions to nonnaturalism are not accidental

---

[114] Actually, one of Plato's central arguments for the introduction of 'the theory of forms' is broadly similar to Moore's 'open question argument' for the introduction of goodness as a nonnatural property. For this see my 'Plato, Moore and Ethical Nonnaturalism' (2009). For the introduction of the theory of forms see also R.Dancy (2004).

because the ontological and epistemological commitments of nonnaturalism are, indeed, far from easy to accommodate. Nonnaturalists, as we mentioned in section 5.2, are aware of this issue and employ a number of different strategies in their endeavour to extenuate their theories from these charges. Some follow Plato and think that we need to vindicate our pre-theoretical realist intuitions no matter the cost and nonnaturalism is the only obvious way to do so. As a result they downplay or even simply ignore these problems or bite the bullet and claim that we are forced to countenance these problems (and the resultant mystery), if we are to make sense of the nature of normative reality[115]. Others again try to meet the discontents head-on in a way that aspires to dispel the clouds of mystery[116].

But let us now sketch what are these infamous ontological and epistemological problems of nonnaturalism. Let us start with the epistemological problem: 'the epistemic access problem'. As the label implies, the problem is that it is not clear how we could ever come to track the existence of nonnatural properties. That is, how we could ever come to identify instances of nonnatural properties. For the ordinary way of coming to know the existence of property instances is through causal interaction of our perceptual mechanisms with the respective instances. We can taste honey and drink water, touch trees, see a bird, smell a rose, hear a concert etc.

These are all instances of natural properties we come to detect and identify through the causal interaction of our perceptual mechanisms with the particular instances. And we can causally interact with these natural properties because they have independent causal powers. That is, they have difference-making powers and they can bring about effects. Water can dissolve salt for example, it can get you wet, conduct electricity, clean your stained shirt etc. and these can be perceived by the mechanisms of our five senses: taste, smell, vision, touch and hearing.

In contrast, this does not seem to be the case with nonnatural properties. Nonnatural properties do not seem to have independent causal powers and therefore are not in any obvious way tracked by our perceptual mechanisms. We don't seem to

---

[115] Arguably, this was Moore's (2000) dialectical strategy.
[116] R.Shafer-Landau (2003) and R.Wedgwood (2007) are two recent normative nonnaturalists that have heroically tried to meet head-on the epistemological and ontological problems of nonnaturalism. Unfortunately, I don't have the space and time to go through their sophisticated views here, although I have done some work on Shafer-Landau's approach. See my 'Causal Exclusion and R.Shafer-Landau's Nonnaturalism' (ms). For a critique of Shafer-Landau's attempt to explain normative supervenience see M.Ridge (2007c).

causally interact with a nonnatural property of justification, goodness or other normative properties often taken to be nonnatural. To paraphrase what Moore (2000:175) once said about goodness, justification (and other epistemic properties) are not properties that we can take up in our hands and meticulously examine. We can't examine them even with the most advanced scientific instruments and devices[117]. Thus, since nonnatural properties do not seem to have any independent causal powers, it is not clear how we could ever track them down and come to know their existence.

Plato himself, as we have seen in the *Phaedo* quote above, talked about nonnatural properties as abstract properties (i.e. 'forms') with the causal power to individuate property instances, no matter how this causal relation of individuation could have been established. Aristotle and other modern philosophers, though, gave and still give Plato a hard time for this causal individuation claim[118]. For, it is not clear how a nonnatural property could ever have independent causal efficacy. That was the problem in the first place. Nonnatural properties do not seem to have independent causal powers and, hence, we don't really know how we could ever come to know the existence of such properties[119].

Here is a nice quote from Aristotle's *Metaphysics* (1079b10-1080a5)) castigating Plato's causal individuation claim:

> 'Above all we might examine the question what on earth the forms contribute to sensible things... for they are not the cause of any motion or change in them. Moreover... [t]o say that the forms are paradigms, and that other things participate in them, is to use empty phrases and poetical metaphors; for what is it that fashions things on the models of the forms? ... Further, it would seem impossible that the essence and the thing of which is the essence exist in separation; hence how can the forms, if they are the essences of things, exist in separation from them? It is stated in the *Phaedo* that the forms are the causes both of existence and of generation. Yet, assuming that the forms exist, still the things which

---

[117] Compare Moore (2000 : 175) : 'It is immediately obvious that when we see a thing to be good, its goodness is not a property which we can take up in our hands, or separate from it even by the most delicate scientific instruments, and transfer to something else.'
[118] See S.Shoemaker (2004) and D.M.Armstrong (1980: 66-68)
[119] This criticism also applies to epistemic kinds realism, if given a nonnaturalist reading.

participate in them are not generated unless '*there is something to impart motion*'; (my emphasis)[120].

Aristotle's point is that to talk of a nonnatural property Fness in virtue of which F things are F is an 'empty phrase and poetical metaphor' because it seems impossible that the Fness property could be distinct from F instances. And it is impossible to be distinct because, if they are distinct, then we have to assume that 'there is something to impart motion'. That is, -in modern terms- something establishing the causal individuation relation between the nonnatural property Fness and F property instances. For the 'in virtue of' that stands between Fness and F property instances seems to be causal. But it seems, Aristotle points out, that 'forms' do not have any independent causal powers, as they don't seem to cause anything in nature. Therefore, it appears implausible that the nonnatural Fness property causally individuates F instances of natural properties.

The traditional response of the nonnaturalist to 'the epistemic access problem' has been to appeal to an 'intuitionistic epistemology'. That is, to propose that such sui generis properties are somehow tracked and identified in a direct and non-inferential way by a cognitive faculty of intuition[121]. This faculty of intuition can somehow directly and non-inferentially reveal to us, 'clearly and distinctly', instances of such properties as if it is a cognitive 'sixth sense' of some sort. It is supposed to somehow function like a cognitive mechanism finely tuned to detect and identify instances of nonnatural properties. Once the faculty of intuition tracks a nonnatural property, then the agent can somehow 'see' the instance of the property. He can 'perceive' that a certain action is good, a belief is justified, a statue is beautiful, a taxation programme is just etc.

But let us narrow down our attention to the object of our discussion: epistemic justification. The idea is that when an agent scrutinizes a certain belief, theory, explanation etc. the faculty of intuition will reveal to him whether the belief, theory etc. is justified or unjustified. It will reveal to him the nonnatural property of justification (or unjustifiedness). This 'revelation' will come in the form of an

---

[120] Translation by H.Tredennick and G.C.Armstong (2006) with some own modifications in order to make it more approximate to modern philosophical parlance.
[121] Philosophers usually considered intuitionists are (middle) Plato, H.Sidgwigk (1967), G.E.Moore (2000). Contemporary philosophers usually considered intuitionists include J.McDowell (1994, 1998) and R.Audi (2007). For the history of the development of intuitionism see T.Baldwin (2006). For a brief introduction to ethical intuitionism see P.Stratton-Lake's (2006).

intuition. The agent will intuit or 'see' whether the belief is justified or not. Of such intuitionistic sort were Moore's (2000) moral epistemology and also the classical and non-classical foundationalist theories of Chisholm (1966) and Bonjour (1998) we encountered in Ch.3. They were appealing to the idea of a cognitive mechanism of 'intellectual seeing', 'self-presentation' or 'rational intuition' that can identify which beliefs are justified and which are not.

Yet the appeal to intuitionistic epistemology does not seem to resolve 'the epistemic access problem'. Rather, it only seems to push it one step back[122]. For it still remains unclear how the faculty of intuition tracks and identifies instances of the nonnatural property of epistemic justification. It remains unclear because nonnatural properties do not seem to have independent causal powers and, therefore, remains enigmatic how any cognitive mechanism could track and identify such properties. For to identify instances of properties we need to causally interact with the relevant properties and, since nonnatural properties do not have independent causal powers, it is not evident how we could ever do so.

The whole idea of a faculty of intuition seems to rely on a vague ocular metaphor of 'seeing' which beliefs are justified and which are not[123]. But the metaphor needs to be cashed out in more exact and explicit terms, if intuitionistic epistemology will dodge the charge of being unduly mysterious. For my part, I don't see how the metaphor could be cashed out in such terms because the metaphorical element seems to be part and parcel of intuitionistic epistemology. If we try to cash out the metaphor in literal terms, then intuitionism seems to collapse. It collapses because, as I said, it remains enigmatic - if this property does not appear to have independent causal powers- how we 'see' the nonnatural property.

It seems then that Kant (2003) was right to say that:

> 'if we understand by it [nonnatural property] an *object* of a *non-sensible intuition*, we thereby presuppose a special mode of intuition, namely, the intellectual, which is not that which we possess, and of which we cannot even comprehend the possibility (B307/p.268)...For the

---

[122] Compare P.Stratton-Lake (2006:6): ' An important factor in the decline of intuitionism was its commitment to metaphysical and epistemological views that were regarded as both deeply implausible and unnecessary'.
[123] See R.Rorty (1979) for how ocular metaphors -largely due to Plato- have dominated western philosophy.

intelligible would require a quite peculiar intuition which we do not possess, and in the absence of this would be *for us* nothing at all (B336/A280/p.288)… for we are acquainted with no kind of intuition but our own sensible kind and no kind of concepts but the categories and neither of these is appropriate to a non-sensible object (A287/p.292)'; (Kant's own emphasis)[124].

Kant's point is that for the tracking of nonnatural properties we would need a cognitive mechanism that goes beyond the boundaries of a naturalistic picture of psychology. That is, we would need a special kind of cognitive mechanism that somewhat mysteriously tracks the existence of nonnatural properties; in Kant's philosophical idiom 'noumena'. But as Kant underlines, *for us* this would be nothing at all, because all we have as human beings is 'sensible intuition' and 'no kind of concepts but the categories'. That is, perceptual mechanisms and conceptual repertoire, namely, a language.

The epistemological problem interlocks with the ontological problem: 'the queerness problem'. The queerness problem owes its name to J.Mackie (1977:38) who claimed that it is not clear what such nonnatural properties could ever be if they existed. If such properties existed they wouldn't resemble in any way the mundane natural properties that impinge on our senses in our everyday life[125]. Natural properties are of the usual ilk. They have independent causal powers and we can easily detect them. Natural properties are difference-making properties and as such they bring causal effects about.

But nonnatural properties do not appear to be like that. They are not like natural properties in any obvious way. If nonnatural properties had independent causal powers we would be able to track them down and identify their instances. But it is exactly because they don't appear to be natural and thereby have any such powers that were postulated as nonnatural in the first place. Since we can't detect causally impotent properties, it is not clear what these nonnatural properties are or even could be.

---

[124] Translated by N.K. Smith. Compare Mackie (1977:38): ' …if we were aware of them, it would have to be by some special faculty of moral perception or intuition, utterly different from our ordinary ways of knowing everything else'.
[125] Compare Mackie (1977:38): '…they would be entities or qualities or relations of a very strange sort, utterly different from anything else in the universe.'

Moreover, it is not only difficult to understand what such nonnatural properties are or even could be but is equally difficult to understand where we should *place* these properties if they exist. Obviously, the natural world won't do because things that exist within the natural realm have independent causal powers and in virtue of these powers, sooner or later, come to our attention. We come to know their instantiation when we perceive relevant instances.

But nonnatural properties are not supposed to be of such kind. They are not supposed to appear within the realm of nature. Otherwise, we would expect them to have independent causal powers that it seems they do not. They do not seem to be embedded in causal relations in any obvious way. Thus, talk of nonnatural properties is shrouded with a ring of enigma and mystery that makes naturalistically minded philosophers to shake their heads in disbelief and perplexity.

This much concludes the brief presentation of the interlocking epistemological and ontological discontents of nonnaturalism. In the next section 5.4 we present the second part of our criticism of nonnaturalism. We examine whether the 'naïve' nonnaturalist can explain epistemic supervenience.

5.4 'Naïve' Nonnaturalism and Epistemic Supervenience

In this section, we present the second part of our criticism of 'naïve' nonnaturalism. We argue that nonnaturalism has problems in explaining 'the epistemic supervenience desideratum'. First, we introduce 'the epistemic supervenience desideratum' and elaborate a bit on *why* it is a constraint on epistemic discourse that any adequate theory of epistemic justification must accommodate. We explicate why it has this binding character on any adequate theory of epistemic justification. Second, we argue that 'naïve' nonnaturalism does not seem capable of adequately explaining 'the epistemic supervenience desideratum'. Let us start with the introduction and elaboration of 'the epistemic supervenience desideratum'.

The 'epistemic supervenience desideratum' is sometimes introduced along with the analogous 'moral supervenience desideratum' that equally constrains moral discourse[126]. This is far from accidental. Both moral and epistemic discourses are normative, rational discourses and, as such, reasons can be asked and be given for

---

[126] See for instance J.Kim (2008). For epistemic supervenience per se see R.Feldman (2001a) and H.Vahid (2005).

our actions and beliefs. Like in moral discourse we are concerned with what to do, in epistemic discourse we are concerned with what to believe. That is, in virtue of a natural property we confer positive epistemic status ('...is justified, well-grounded, rational etc.) to beliefs and -all other equal- we form the beliefs in question.

As in moral discourse we make moral evaluations that, all other equal, lead to action, in epistemic discourse we make epistemic evaluations that, all other equal, culminate in the fixation of belief. But our moral and epistemic evaluations are not just ad hoc or arbitrary. They are *based on* reason-giving natural properties (even if these reasons are grossly capricious or plainly irrational). We form beliefs because we find them coherent, rationally intuitive, because they are motivated by epistemic virtues, because the informant is trustworthy, or even because our 'mum says so' etc. It is this *'reason-ableness'* that makes supervenience a binding constraint on epistemic discourse and, therefore, a desideratum calling for adequate explanation.

For, we evaluate an event, action etc. as good and a belief as justified, reasonable etc. because we think that there is a certain natural property giving, respectively, a reason for action or a reason for belief. In other words, the normative properties (goodness, justification etc.) that guide our moral and epistemic conduct supervene on non-normative, natural properties. This imposes a *rationality constraint* on moral and epistemic discourses. Any theory of epistemic justification should explain how this tie of our epistemic evaluations on natural properties takes place. How normative properties are anchored on natural properties.

This rationality constraint on epistemic discourse implies that the very same natural property cannot realize contradictory epistemic evaluations, though, the converse does not hold. Namely, there can be the same epistemic evaluation supervening on or being multiply realized by different natural properties. In short, there can be no two contradictory epistemic evaluations of two numerically identical states of affairs, events etc., unless these state of affairs etc. have some difference in terms of natural properties (though, the converse does not hold).

Supervenience theses are usually given a *global* reading in terms of possible worlds[127]. According to global supervenience, two entire possible worlds cannot

---

[127] See M.Ridge (2003) and B.McLaughlin and K.Bennett (2005). Also, possible worlds talk should not be taken to have the ontological implications that notoriously D.Lewis (1986) took them to have. Like S.Kripke (1981) I take possible worlds talk to be only a useful conceptual tool and remain neutral about the ontological implications. Besides, there are also antirealist understandings of

instantiate different normative properties, unless they also differ in at least some of their natural properties (but not vice versa). To deny this thesis is to accept that there could be two numerically identical possible worlds in terms of natural properties, but differing in the supervening epistemic properties these worlds instantiate.

For example, we could have two numerically identical possible worlds with the same natural properties (geological data like fossils etc.) and the first possible world instantiates epistemic justification for the belief that earth is many millions years old, while the second world instantiates epistemic unjustifiedness. This would surely be bizarre.

To fully grasp the vital importance of the supervenience relation on normative discourse, let us for the sake of the argument grant the counterfactual possible world were the supervenience relation on normative discourse does not hold. This scenario has some profoundly bitter repercussions. Repercussions difficult to be exaggerated, strongly suggesting that such a 'supervenience-free' counterfactual possible world is not really, metaphysically possible.

For, if the supervenience relation between normative properties and non-normative, natural properties was possible somehow to be severed, then it would seem that the rationality constraint on normative discourse just collapses. The upshot of the collapse of the rationality constraint would be that our moral and epistemic evaluations are not really based on reasons. They don't supervene on certain reason-giving natural properties. They are not anchored on natural properties but they are somehow 'free-floating'. As one might surmise, this is as bad as it could get.

For, as we have said, the rationality constraint implies that the very same natural property cannot realize different epistemic properties. If the very same natural property could realize different epistemic properties, then epistemic discourse would simply cease to be a rational discourse. That is, a discourse for which the reasons grounding our epistemic evaluations can be asked and be given. It would cease to be a field of inquiry where there can be a legitimate exchange of the reasons supporting our epistemic evaluations.

It would cease to be a rational discourse because asking and giving the reasons supporting our epistemic evaluations would prove just to be 'idle talk'. Two agents

possible worlds talk. See for example S.Blackburn (1993: 52-75).

124

could ground their contradictory epistemic evaluations on the very same natural property and, hence, any rational examination of the reasons supporting the epistemic evaluation would have no dialectical force, no dialectical implications. It would just be rhetorical idle talk with no potential of ever reaching convergence and that would be the end of the rational practice of asking and giving reasons for our epistemic evaluations.

This bitter repercussion is far-reaching and, I think, should be considered to be *a reductio ad absurdum* against anyone who might be tempted to think that we can reject without impunity the moral and epistemic supervenience desideratum. It shows that we should better take the supervenience of the normative on the natural very seriously, if we don't want to taste this bitter repercussion.

However, despite the intuitive plausibility of the epistemic supervenience desideratum, a challenge to the desideratum might come from strong forms of epistemic externalism. Such externalists might stress that the epistemic supervenience desideratum rests on the internalist 'accessibility requirement' and, once this is discarded, the supervenience desideratum is sidelined. It is not any more a desideratum.

This thought might be prompted by the fact that the epistemic supervenience desideratum is set in terms of an epistemic property supervening on a natural property and this seems to imply that the agent needs to have direct access to the reasons supporting his epistemic evaluation. Yet the externalist denies that we need to have direct access to the reasons supporting our epistemic evaluations. We can have justified beliefs without being aware of the reasons or grounds supporting these justified beliefs. Hence, the supervenience constraint is not a much of desideratum for the undaunted externalist[128].

This point is, I think, well-taken by the externalist. But this should not really create any trouble for us because, as we clarified in section 1.4, we don't ally ourselves with such a form of strong externalism, though, we don't ally ourselves with a strong form of internalism either. As the reader might recall, our stance has been reconciliatory towards the internalism/externalism controversy. We didn't expand the exact details of our reconciliatory stance in that context and the current

---

[128] A.Goldman (1991) and T.Williamson (2000) might be such externalists.

context remains equally inopportune for such a task. The exact detailed explication should wait for a more convenient circumstance. But for current purposes our short discussion is, I think, sufficient. Since we don't commit ourselves to strong externalism, we should feel no challenge coming from these externalist considerations against the supervenience desideratum.

To wrap up the above discussion, we have found the epistemic supervenience desideratum to be plausible. On the one hand, renouncing the epistemic supervenience constraint seems to lead to a reductio and, on the other hand, externalist considerations do not really touch us. Reasonably then, the supervenience of normative properties on non-normative, natural properties has been considered so plausible that many philosophers would concur and consider the supervenience relation *necessary* (and a conceptual truth). That is, necessarily, there can be no possible world where normative discourse is not constrained by the supervenience relation.

With the epistemic supervenience desideratum introduced and motivated, the promised second step is to pose the question whether 'naïve' nonnaturalism does explain the desideratum. The answer seems profound enough. It seems that it does not. Let me explain. An adequate explanation of the epistemic supervenience desideratum demands an account of how epistemic properties supervene on non-epistemic, natural properties. That is, how epistemic properties are being grounded or anchored on natural properties.

According to the global reading we gave to epistemic supervenience, two entire possible worlds cannot differ in their epistemic properties, unless they also differ in their natural properties (but not vice versa). To deny this thesis is to accept that there could be two numerically identical possible worlds in terms of natural properties, but differing in the supervening epistemic evaluation about these worlds. This would be absurd.

The nonnaturalist, presumably, as a response to pressure to explain epistemic supervenience, would be forced to concede the following account. The property of epistemic justification is nonnatural and supervenes on a certain natural property or a certain conjunctive concatenation of natural properties. For example, epistemic justification could supervene on coherence or self-presentation or the conjunction of

coherence and empirical evidence etc.

Yet, such an account seems entirely hopeless. For, it is clearly possible that the nonnatural property of epistemic justification in at least some cases might not supervene on any natural property we might think of. This is clearly possible because we can't see how to forge a metaphysical supervenience relation between two different realms of properties: one natural and one nonnatural. For example, let us for the sake of argument say that the nonnatural property of epistemic justification supervenes on coherence. Then it seems clearly possible that there might be coherent beliefs, theories etc., which are not justified and do not realize epistemic justification. In terms of possible worlds, it seems clearly possible that we can have two identical possible worlds in terms of natural properties and still have a different epistemic property realized by the two worlds.

This much seems to follow from 'the Moorean/Humean lesson' because it remains an 'open question' whether epistemic justification is reducible to any natural property. If epistemic justification was reducible to a certain natural property, then 'the Moorean/Humean lesson' would have been blunted and epistemic justification would trivially supervene on the natural property reducing epistemic justification. For, trivially, any property supervenes on itself. But as we have argued, and the nonnaturalist agrees, this is not what happens. 'The Moorean/Humean lesson' holds fast and seriously undermines such reductive efforts. Thus, the nonnaturalists' attempt to explain supervenience seems doomed from the very beginning.

As a result, I conclude that 'naïve' nonnaturalism does not seem able to adequately explain the epistemic supervenience desideratum and this is surely a serious liability. It is a serious liability because as we have discussed, any normative discourse is constrained by a supervenience relation. Any normative discourse is a rational one and as such reasons can be asked and be given for our beliefs. And if in situations identical in terms of natural properties we have contradictory epistemic evaluations then the discourse is not rational any more. Rational exchange of argument collapses and epistemic talk proves to be just idle talk. That is, talk that has no impact on the way we form, retain and revise our beliefs. Belief-fixation, retention and revision turns out to be simply capricious.

Overall, the epistemological and ontological discontents coupled with the

apparent difficulties nonnaturalism has with the epistemic supervenience desideratum seem to prompt the conclusion that there is no nonnatural property of epistemic justification. In terms of 'the epistemic justification puzzle', this is translated as 'the nonnaturalism-denial assertion' (P4) is plausible: justification assertions and attributions do not purport to refer to a nonnatural property of epistemic justification.

But this inference might be too quick, as J.McDowell has defended a subtle and sophisticated nonnaturalist theory in a naturalistic cloth. In the next section we examine McDowell's interesting proposal.

5.5 J.McDowell's Quietist Nonnaturalism Introduced

J.McDowell (1994, 1998) has pursued a quietist line in order to rescue normative realism[129]. McDowell is not any typical sort of 'naïve' nonnaturalist. In fact, he (1994:77-8) renounces what he calls 'rampant Platonism' and declares himself a committed naturalist. Albeit, a naturalist of a neoaristotelian sort: 'naturalist of second nature'. That is, a naturalist that steers between the straits of extravagant 'rampant Platonism' and narrow-minded scientistic (*not* scientific) 'bald naturalism'. Let me explain because this is not a mere play of the word 'naturalism'.

Inspired by Aristotle, McDowell (1994: xix-xxiv, 1998) distinguishes between two forms of naturalism: 'bald naturalism' and 'naturalism of second nature'. Bald naturalism is the naturalism that equates nature with what he calls 'the realm of law'. That is, with the realm of independent causal efficacy constantly being explored by scientific inquiry. This is a fairly standard way of understanding the term of art 'naturalism' and this is how we ourselves understood naturalism and natural property in Section 2.2.

'Naturalism of second nature', however, does not identify nature with 'the realm of law'. To do so, McDowell suggests, would have been to submit to the unwarranted scientism of 'bald naturalism' which mistakenly identifies nature with 'the (causal) realm of law'. Historically speaking, McDowell (1994:70-72) continues, this biased equation of nature with 'bald naturalism' and 'the realm of law' has specific traceable origins. It is a modern construction that came about under the impact of the rise of modern science (Galilean revolution etc.) and its impressive

---

[129] A similar quietist approach to McDowell's has been sketched by P.Strawson (1987).

achievements. Nothing of it exists in Aristotle's pre-modern conception of naturalism and it is to such (uncontaminated by the impact of science) conceptions of naturalism that we need to return.

Indeed, McDowell's own positive proposal follows the dictates of this diagnosis and the plea for a return to an uncontaminated Aristotelian form of naturalism. In contrast with 'bald naturalism', McDowell's own neoaristotelian 'naturalism of second nature' leaves open logical space for what he calls (after W.Sellars (1997)) 'the logical space of reasons'. That is, the precious realm of (human) rationality anchored, as Sellars (1963:39-40) would have put it, in 'the framework of being a person'.

According to McDowell's neoaristotelian virtue-based account, a rational agent is one that, if being virtuously brought up in a social context, responsiveness to the right practical reasons for action and the right theoretical reasons for belief will come to be his 'second nature'. He will come to acquire a practical know-how (Aristotelian 'phronesis' or practical wisdom) of what to do and what to believe from case to case and from context to context[130]. Yet, 'the space of reasons' is not one that can be included in 'the realm of law'[131]. Reasons are one thing, causes are another and their respective domains of explanatory responsibility are clearly distinct and should not be conflated.

Here is a nice quote (1994:82-84):

'… a decent upbringing initiates us into the relevant way of thinking, our eyes are opened to the very existence of this tract of the space of reasons… ordinary upbringing can shape the actions and thoughts of human beings in a way that brings these demands into view…The point is clearly not restricted to ethics. Moulding ethical character, which includes imposing a specific shape on the practical intellect, is a particular case of a general phenomenon: initiation into conceptual capacities, which include responsiveness to other rational demands besides those of ethics. Such initiation is a normal part of what it is for a human being to come to maturity… one's eyes [are] opened to reasons at large by acquiring a second nature'.

---

[130] A full-blown virtue epistemology of this broad ilk has been developed by L.Zagzebski (1996).
[131] In D.Davidson's (2003: 42) terms, 'the constitutive ideal of rationality'.

129

Armed with the 'bald naturalism'/ 'naturalism of second nature' distinction, McDowell (1994) then insists that, by any means, we should not be allured by 'bald naturalism' which identifies nature with 'the realm of law' and promises, in Max Weber's (1982) famous expression, to fully 'disenchant the world'[132]. That is, promises in the prospect of a mature science to bring natural workings under nomological causal regulation. Nature cannot be fully disenchanted as the enlightenment project has arrogantly promised; or at least human nature. Human nature resists disenchantment because the realm of reasons, which is the key mark of being a human primate, eludes inclusion to the realm of causes.

For human nature decisively belongs to the realm of reasons, not causes. We explain and predict intentional action by appeal to reasons for action, we explain and predict belief-fixation by appeal to reasons for belief etc. and not by any appeal to causes[133]. Any appeal to causes for an explanation of such phenomena seems hopelessly irrelevant. And for McDowell, this is *a brute fact* about the constitution of the natural world. And it is here that the quietist element enters the scene. For this is simply how things are and we cannot further analyze why this is the case. This is how far philosophy can go in terms of addressing the core philosophical puzzles (normativity, mental causation, free will, meaning, personal identity etc.). Philosophy, thus, rests in peace only when this line is drawn (and realized) between what we can understand and what we cannot understand.

To come to our specific point of interest, McDowell (1994) argues that once we grasp this fine-grained distinction between 'bald naturalism'/'naturalism of second nature', the concomitant renunciation of 'rampant Platonism' and that this theory is only a modest 'naturalized Platonism', the epistemological and ontological discontents evaporate. They evaporate because we neither postulate the existence of any 'queer' Platonic, nonnatural properties nor of a spooky faculty of intuition.

Rather, we only make the modest claim that as rational animals (initiated through acculturation into the social realm), we come to acquire the cognitive ability to be responsive to reasons (practical and theoretical). That is, through initiation to the

---

[132] Compare M.Weber (1982): '…rational, empirical knowledge has consistently worked through to the disenchantment of the world and its transformation into a causal mechanism' (1982:350) and 'The fate of our times is characterized by rationalization and, above all, by the disenchantment of the world' (1982:155).

[133] See M.Smith (1994), G.E.M. Anscombe (2000), D.Davidson (2001b) and D.Dennett (2003) for discussion of action explanation by appeal to reasons.

social realm we actualize 'the structure of the logical space of reasons'. This potency for acquisition of the cognitive ability for reasons-responsiveness constitutes part of our natural endowment and it should be taken as primitive. It is simply a brute fact about the natural constitution of our cognitive architecture. Nothing more informative can be said about 'the logical space of reasons' and the philosophical puzzles associated with it[134].

This concludes our brief presentation of McDowell's quietist proposal. Let us now turn to the critical examination of his proposal.

5.6 J.McDowell's Quietist Nonnaturalism Criticized

While McDowell's heroic proposal has been well received by some philosophers, other philosophers have expressed strong reservations[135]. For McDowell's proposal has been criticized as only sweeping the epistemological and ontological discontents under the carpet of quietism, though, not really answering them. And this seems fair enough to my eyes. For it does not explain neither to *what* normative properties we are sensitive and responsive to as rational animals nor *how* we are sensitive and responsive to such properties, if these properties do not populate 'the realm of law'. That is, how we are sensitive to an action's property of goodness or a belief's property of justification and what kind of properties these normative properties are, if these properties do not reside in nature and have independent causal powers.

These were basically the 'queerness' and 'epistemic access' problems of nonnaturalism we encountered in the last chapter and McDowell's theory does not seem to really address them. As far as I can see, McDowell's claim that as rational animals we have the cognitive ability to be responsive to reasons (practical and theoretical) is not all that modest, unless we are able to spell out in lucid naturalistic terms how this reasons-responsiveness ability works. That is, what is t*he modus operandi* of 'the logical space of reasons'. How it tracks normative properties in spite of the fact that these do not populate 'the realm of law'.

McDowell himself seems to be aware of this line of criticism but thinks he can rebut it. As he (1994:82-3) says:

'…the very idea of sensitivity to real demands of reason looks spooky,

---

[134] It should be noted that McDowell (1994: 123-4) appears hostile to evolutionaty considerations that might be called in to help us out with accounting for 'the logical space of reasons'.
[135] See for example P.Goldie (2002) who seems to follow McDowell.

unless we can reconstruct it from materials that are naturalistic in the relevant sense' and '…the outlook I am attributing to Aristotle looks like a kind of Platonism. But it is not what I called 'rampant platonism'. We fall into rampant Platonism if we take it that the structure of the space of reasons is *sui generis*, but leave in place the equation of nature with the realm of law. That makes our capacity to respond to reasons look like an occult power, something extra to our being the kind of animals we are, which is our situation in nature' (McDowell's own emphasis).

Yet, unlike McDowell, I don't think he can rebut this line of criticism just by recourse to neoAristotelian 'naturalism of second nature'. And he can't because talk of 'naturalism of second nature' is not to 'reconstruct the very idea of sensitivity to real demands of reason from materials that are naturalistic'. These 'materials' are naturalistic only in McDowell's 'relevant sense' of naturalism and this 'relevant sense' stretches naturalism to a breaking point. It stretches naturalism to the point of being nonnaturalism in naturalistic (in McDowell's 'relevant sense') disguise. And, unfortunately, this is only to push the problem one step back but not to really rebut it. Let me explain why I think this is the case.

The worry I have in mind has been well made by J.Fodor (1998). J.Fodor (1998) chastises the fact that all McDowell has to provide in relation to these problems is picturesque ocular metaphors and that his 'naturalism of second nature' is not, after all, naturalistic itself. Given that Fodor (1998:7-8) expresses the point nicely, I quote freely:

'Having situated the rational (and the ethical, and a lot else that we care about) outside the realm of law, McDowell needs to face the embarrassing question how, by any natural process, do we ever manage to get at it?... Likewise, and more so, in the nonperceptual cases, where the objects of cognition are normative or otherwise intentional aspects of things: how do we get at those if they aren't in the natural order? Maybe better, how do they get at us? *How can what is not in the realm of law make anything happen*? … 'Bringing into view' is a metaphor; only what is in nature can literally be viewed. And 'resonating' is also a metaphor; only what is in nature can be literally attuned to ... The forms of human sentience resonate, as far as

anybody knows, only to aspects of the 'disenchanted' world. Mere exhortation won't fix that. Second nature is what we get when 'our *Bildung* actualizes some of the potentialities we are born with;

… But the question arises how second nature, so conceived, could itself be natural. It is no good for McDowell just to say that it is and you can get some at the *Bildung* store; he has to say how it could be short of spooks… a dualistic naturalism isn't in the cards. If that's right, then epistemology needs to bend and McDowell will have to cool it a little about justification. Justification can't require what can't happen, on pain of there not being any; and whatever happens, happens in the realm of law'; (Fodor's own emphasis).

As we already saw in our discussion of intuitionism in section 5.3, ocular metaphors are part and parcel of the problem and, hence, Fodor is right in his criticism of McDowell. Ocular metaphors cannot constitute a real response to 'the epistemic access' and 'queerness' problems. And this is not simply to inveigh against any philosophical use of metaphors. Metaphors can be useful and suggestive philosophical tools as far as they are harmless and innocuous in terms of implications[136].

McDowell's use of ocular metaphors, however, is not so harmless because it obscures how we are sensitive to normative properties (if at all) and what kind of normative properties we are sensitive to (if any). That is, how we are sensitive to normative properties that do not reside in the causal realm of nature, McDowell's 'realm of law'. To normative properties that do not seem to have independent causal powers and, therefore, are not in any obvious way difference-making properties. They don't bring about any change in the realm of nature.

McDowell's argument aside, it seems that his 'naturalism of second nature' turns out to be, in essence, only a variant of nonnaturalism in naturalistic disguise. It seems to be a variant of nonnaturalism in disguise because, although denounces the 'rampant Platonism' of nonnatural properties, he still speaks in terms of the language of ocular metaphors we chastised in section 5.3. Namely, in terms of 'seeing' that a

---

[136] Of course, metaphors are tricky tools and have to be used with caution in philosophical debates. For one thing metaphors are often ambiguous and susceptible to multiple interpretations and this makes it difficult to grasp their full inferential implications. See W.Lycan (2006: Ch.14) for discussion of metaphor.

certain situation merits a certain response, of 'being sensitive' to normative properties, of 'bringing into view' etc. The language of ocular metaphors, though, seems to be nothing more than the vehicle of a subtle, covert articulation of nonaturalist positions; unless of course they are spelled out in naturalistic terms something McDowell does not perform.

McDowell still has a subtle rejoinder, though. His rejoinder would probably be that with our 'equation of nature with the realm of law' we unduly restrict the domain of nature and slide back to the snare of bald naturalism's philistine scientism. A snare McDowell has taken pains to warn us about. That is, a scientism that does not leave open logical space for the reality of such precious things like norms of rationality, normative facts and truths, mental states etc. and aspires to fully 'disenchant the world'.

For such entities, if exist at all, they don't seem to exist in nature, in 'the realm of law'. And if such precious entities are to be excised, then surely something has gone badly wrong in our contentions. The obvious candidate of what has gone badly wrong is our commitment to 'bald naturalism' and, therefore, we should discard it and embrace his own more liberal 'naturalism of second nature' that does not fall prey to the excision of philosophically precious notions.

My line of response is Sellarsian (and indeed I think Sellars' (1963) own attitude)[137]. We need first to distinguish between being committed to bald naturalism and being committed to scientism. The two claims should not be conflated so light-heartedly because bald naturalism need not entail scientism, (though, the reverse does hold). McDowell, however, conflates the two commitments while I think (with Sellars) we should not. In Sellars' (1963:32) terms, one can be committed to 'the primacy of the scientific image' over 'the manifest image of perennial philosophy' and still resist scientism. This might appear surprising to some people, but I think it should not. Let me explain.

Although one can be committed to 'the primacy of the scientific image' and the verdicts of 'the realm of law', he might still remain firmly pessimistic about the possibility of 'the space of reasons' (and related notions like values, norms, truths,

---

[137] Although I can't flatter myself for being a Sellars' scholar, it seems to my eyes that this has been Sellars' position as well. But for more on this interpretation I should better point to Sellars' sholars. See W.de Vries (2005) and J.O'Shea (2008).

mental states etc.) being reduced to naturalistic terms. Like Sellars, although I am of course committed to the primacy of a scientific-naturalistic worldview, I remain deeply sceptical about the reduction (not to mention elimination) of 'the space of reasons' to nomological causal relations. The mental is and will remain, I think, 'anomalous' in D.Davidson's (2003) memorable characterization[138].

I remain sceptical about the possibility of being fully naturalized, as eliminativist materialists like P.Churchland (2003) think will happen sometime in the future with the maturation of neuroscience. Of course, this deep scepticism does not, naively, preclude the logical possibility of such an event, which clearly appears to be wide open. I am only sceptical about the metaphysical possibility of such an event because the step from logical to metaphysical possibility (as we know from modal arguments in philosophy of mind) is a big one and I really doubt whether this step could ever be taken.

Why I remain sceptical about the metaphysical possibility of such a reductionism is a big question to pursue here, but I can afford a general hint. In principle, my thorough scepticism is not motivated by any quaint feeling that the mental is sacrosanct, inviolable etc. or any other such backward form of thinking. This wouldn't comport with my commitment to a scientific-naturalistic picture. I am a pessimist about the metaphysical possibility of such a reduction because I don't quite see how this is to be done either on a priori conceptual grounds or by empirical means at all.

Analytic naturalistic reductionism as an approach to normative concepts employed by 'the space of reasons' (rationality etc.) fails since Plato and neuroscience despite its advancements does not seem able to reduce the mental (propositional attitudes or qualia) to neural activity[139]. On this point, for reasons that I cannot belabour more here, my Sellarsian intuitions incline me to side in agreement with McDowell (1994) and Davidson (2003).

This thorough pessimism is also wed with a certain -Sellarsian again- background picture about the relation between philosophy and the special sciences. This

---

[138] Davidson (2003) makes clear that he is talking only of propositional attitudes and not qualia that appear to be even more difficult to accommodate in a naturalistic context.
[139] The type-identity reductionism of the 50's-60's in philosophy has been famously twharted by Putnam's (2003) 'multiple realizability argument' and ushered theorists towards functionalist approaches like J.Fodor's (1975), D.Chalmers'(1996) and J.Kim's (2000, 2005).

background picture contends that philosophy is not reducible to the special sciences but, in tandem, is not totally independent either. As Sellars (1997:80) have put it, we should not 'confuse the sound idea that philosophy is not science with the mistaken idea that philosophy is independent of science'. Philosophy is autonomous, though, empirically informed; or at least it should be.

As a consequence, philosophical methodology should not be cornered between the unpalatable horns of the dilemma of austere a priori conceptual analysis or abject surrender of the philosophical branches to the special sciences[140]. That is, to wit, between pensive Platonic beard scratching and Quinean naturalization-liquidation of philosophy (e.g. epistemology reduced to cognitive psychology and sociology of knowledge etc.). Science might be very important for philosophy but it is not in the Quinean sense 'the final arbiter of truth' in matters philosophical. The dilemma is not exclusive and a third 'middle' way along the Sellarsian lines we just draw is both possible and plausible[141].

I conclude that McDowell's equation of bald naturalism with scientism is too quick. As I explained following my reading of Sellars (1963), one can legitimately be committed to 'the primacy of the scientific image' and still remain non-scientistic, though, properly scientific. He can do so because remains sceptical about the metaphysical possibility of a naturalization of 'the space of reasons' and the concomitant handing over of philosophy (and its problems) to special scientists (neuroscientists, psychologists, sociologists etc.).

In virtue of this scepticism, I find the bright optimism of philosophers like Churchland and Quine unwarranted, too impressed and allured by the impact of science and its achievements. This much I found indeed scientistic, not merely scientific. Of course, they could also charge me with analogous comments for my- despite the impact and achievements of science- Sellarsian pessimism but given my priorities at hand, I think I have to cut the chase of this issue at this juncture. This fascinating discussion can wait for another occasion.

Last but not least, McDowell does not really take into consideration 'the epistemic supervenience desideratum'. As far as I know, he nowhere addresses the

---

[140] See W.Quine (1953, 2008) and for discussion R.Feldman (2001). Feldman (2001a) calls the Quinean approach 'replacement naturalism'.
[141] This stance is what R.Feldman (2001a) calls 'cooperative naturalism'.

supervenience relation of the normative on the natural. Still, McDowell needs to explain how epistemic properties supervene on non-epistemic, natural properties. Granted McDowell's 'naturalism of second nature' (which we found to be only nonnaturalism in disguise) I don't see how this is to be done.

Our criticism of nonnaturalism in section 5.4 is, therefore, revived. For it is always possible for a nonatural, epistemic property to fail to supervene on a nonepistemic, natural property. That is, no matter which natural property we may invoke it seems that is always possible that the nonnatural, epistemic property might fail to supervene. And this is sufficient to constitute a severance of the supervenience relation of the normative on the natural. It is sufficient to make clear that no matter which natural property we might invoke, the epistemic property might fail to supervene (or be 'anchored') on it.

To sum up, McDowell's 'naturalized platonism' fails to neutralize the problems of nonnaturalism because it does not really address the 'queerness' and 'epistemic access' problems. His distinction between 'bald naturalism'/ 'naturalism of second nature' and the appeal to our primitive natural endowment of 'the space of reasons' (actualized by socialization) does not make the trick. He still needs to explain how we are sensitive to normative properties (if at all) and what these normative properties are (if any), if these properties do not reside in the natural realm.

The playful wrap in the language of ocular metaphors does not suffice to conceal these problems; unless, of course, McDowell spells out the metaphors in lucid naturalistic terms something he does not do (and I do not see how it could be done, at any rate). McDowell also does not address epistemic supervenience and, even if he did, his 'naturalized platonism' does not really seem in position to adequately account for it.

With our criticism of McDowell's quietist proposal, our criticism of nonnaturalism comes to an end. McDowell's quietist proposal has been found inadequate to rescue nonnaturalism from the thorny problems we have highlighted in sections 5.3-5.4.

5.7 Conclusion and Summary of the Argument

In this chapter, I have performed two basic things. First, in section 5.2 I introduced the theoretical commitments of 'naïve' nonnaturalism and the twofold

motivation for pursuing the nonnaturalist project. That is, on the one hand, the accommodation of 'the Moorean/Humean lesson' and 'the epistemic twin earth lesson' and, on the other hand, the explanation of our pre-theoretical realist intuitions about epistemic discourse.

Second, in sections 5.3-5.4 I explained why the 'naïve' nonnaturalist approach seems to be a theoretical giant with feet of clay. Why although it prima facie seems to be particularly auspicious, it threatens to collapse to rubble once we realize its epistemological and ontological discontents. That is, 'the epistemic access' and 'queerness' problems presented in section 5.3 and the apparent inability of explaining the 'epistemic supervenience desideratum' presented in section 5.4. In sections 5.5-5.6 I then looked into McDowell's 'naturalism of second nature' and found his subtle version of nonnaturalism also wanting. Overall, the discussion of the problems of nonaturalism seems to prompt the conclusion that 'the nonnaturalism-denial assertion' (P4) is plausible. Our justification assertions and attributions do not purport to refer to a certain nonnatural property of epistemic justification.

Chapter 5 also marks the end of the critical 'Part I: The Problems of Epistemic Referentialism'. In the next chapter 6, we inaugurate the more constructive 'Part II: The Possibility of Epistemic Expressivism'. We look into a very different approach to normative discourses from the ones we have examined and found implausible so far. Unlike the approaches we have examined and found implausible so far, this approach is nonreferential. That is, it rejects 'the referential semantics assumption' (P2), namely, that 'justification assertions and attributions purport to refer to a certain property (natural or natural) of epistemic justification'. This approach is a nonreferential, expressivist one that comports nicely with the premises of the puzzle we have found to be plausible, namely, 'the naturalism-denial assertion' (P3) and 'the nonnaturalism-denial assertion' (P4).

Part II The Possibility of Epistemic Expressivism

Chapter 6 The Advent of Normative Expressivism

6.1 Introduction

So far, in the previous chapters 2-5 we have criticized referentialist/realist approaches: analytic and synthetic naturalistic reductionism, epistemic kinds realism, 'naïve' nonnaturalism and McDowell's 'naturalism of second-nature'. All this critical work has bolstered the assumption that premises 3 and 4 of 'the epistemic justification puzzle' are plausible, that is, the naturalism-denial and the nonnaturalism-denial assertions are plausible. As these claims state, justification assertions and attributions do not purport to refer to a certain natural or nonnatural property of epistemic justification.

Yet, given that premises 3 and 4 have been found to be plausible and premise 1, namely, 'the property assumption' (stating that all properties are either natural or nonnatural) seems to be a mere truism then, by elimination, a solution to the puzzle should go through a rejection of the remaining premise 2. That is, it should go through a rejection of 'the referential semantics assumption', namely, that justification assertions and attributions purport to refer to a certain property of epistemic justification (natural or nonnatural).

Thus, the argument of the thesis can be seen as *an argument by elimination*. It eliminates one-by-one possible ways out of the puzzle, by means of arguments for their implausibility, and then suggests that the way out should go through the rejection of the remaining 'referential semantics assumption'. At least, it should go down this path if two assumptions hold: the assumption that the arguments we gave for the implausibility of the other ways out do hold and the assumption that such a *non*referential approach will not have any unpalatable implications that could count as a reductio against it. The former assumption is one that we have already embraced and the latter is one that can be evaluated only after we fully explore a nonreferential approach and we haven't yet even started doing this. Hence, it is premature to say

anything about the latter assumption.

Such a nonreferential way out of the puzzle I intend to explore to some depth in this chapter 6 and, more substantially, the next chapter 7. My intention is to reject 'the referential semantics assumption' and explore an expressivist solution to the puzzle. This nonreferential solution is to be an expressivist one because, as I will argue in section 6.4, expressivism easily evades the semantic problems of the realist approaches we have made explicit in chs.2-5 and this *motivates* the possibility of expressivism. At least, it motivates the conceptual exploration of this interesting approach in good hope that it will both prove explanatorily fruitful and free of any counter-intuitive implications.

In particular, expressivism accounts for the Moorean/Humean 'open feel' semantic intuitions as these have been witnessed in 'the Moorean/Humean lesson' and 'the epistemic twin earth lesson', it avoids 'the epistemic access' and 'queerness' problems of the nonnaturalist and promises to account for 'the epistemic supervenience desideratum' that also nonnaturalism has failed to accommodate.

My dialectical intention now made explicit, this chapter 6 is mostly preparatory for the more constructive chapter 7 where I explore an original form of epistemic expressivism. This chapter serves a twofold preparatory purpose: to *introduce* the basic ideas of expressivism as a semantic programme and *motivate* the possibility of such an expressivist approach to epistemic justification discourse by means of showing how it can easily evade the referentialist's problems (just enumerated above). For the mere sake of this twofold purpose, I will use as a 'toy' expressivist theory a theory largely inspired by early A.Gibbard's (1990) influential norm-expressivism[142].

This is then how this preparatory chapter will unfold. In section 6.2 I introduce the semantic programme of expressivism and explain some of its basic ideas. In section 6.3, I outline Gibbard's norm-expressivism about moral rationality and then in section 6.4 explain how it could be applied to epistemic justification discourse (as has already been in a relativist form by H.Field (2000, 2009)). In section 6.5, I explain how our 'toy' expressivist theory of norm-expressivism can easily evade the problems of the various realist approaches we have encountered in chs. 2-5 and how

---

[142] Philosophers influenced by A.Gibbard (1990) include H.Field (2000, 2009), M.Chrisman (2007) and K.Kappel (2010, ms).

this motivates the possibility of an expressivist approach to epistemic justification discourse. Finally, in section 6.6 I preview the argument of this preparatory chapter and pave the way for the more constructive chapter 7.

6.2 What Expressivism Is: Some Basics

In this section 6.2 we introduce the basic ideas of the semantic programme of expressivism. Expressivism is an approach that goes against standard orthodoxy in semantics and its origins can be traced at least back to insights found in the work of Locke and Hume[143]. But let us start with an outline of what standard approaches to semantics suggest and then introduce an outline of the semantic programme of expressivism and how it aspires to contest this semantic picture and establish itself as a full-blown semantic theory.

From Frege onwards, standard semantic theories have been referential and truth-conditional in character[144]. They follow a broadly *referential semantic strategy* and assume that *meaning is reference*. That is, they purport to explain the meaning of sentences in terms of what the subject-predicate of the sentence purport to refer to. That is, if the subject refers to an individual (e.g. grass) that falls within the extension of the predicate (e.g. green), then the sentence satisfies its truth-condition and it is true (i.e. 'Grass is green' is true if and only if Grass is green). If the subject refers to an individual (e.g. my dog) that does not fall within the extension of the predicate (e.g. green), then the sentence does not satisfy its truth-condition and is false (e.g. 'My dog is not green' is true if and only if My dog is not green).

Atomic sentences then like {P,Q}, courtesy of logical operators (conjunctions, conditionals etc.), can be conjoined in logically complex ways like {PvQ} and appear embedded in truth-functional contexts like disjunctive syllogisms {PVQ, ¬P, Q}, modus ponens {P→Q, P, Q} etc. When embedded in such truth-functional contexts, we can easily predict their logical properties (validity, consequence etc.) recursively by reference to the truth-tables of the logical operators of propositional logic and the various rules of inference. In a parallel way we can predict the logical properties of sentences as they are being expressed in predicate logic and beyond. In this sense, in standard truth-conditional semantics linguistic content explains mental

---

[143] See J.Locke's (1973) ruminations on language and D.Hume's (1986) account of moral language.
[144] For discussion of Frege's views and contribution to philosophy of language see D.Wiggins (1999) and C.Travis (2006).

content. That is, language explains thought.

This core semantic idea has been exemplified in 'the referential semantics assumption' (P2) itself. Justification assertions and attributions are true if their atomic parts (subject-predicate) do refer and thereby satisfy their truth-conditions. For example, 'Mike's belief that is raining is justified' is true if and only if the subject, that is, 'Mike's belief that is raining' falls under the extension of the predicate, that is, the property of 'being justified'. Analogously, 'Mike's belief that is raining is justified' is false if and only if the subject, that is, 'Mike's belief that is raining' does not fall under the extension of the predicate, that is, the property of 'being justified'.

Yet, unlike standard truth-conditional semantics that assume a referential semantic strategy, expressivist semantics assume a *use semantic strategy*. A use semantic strategy assumes that *meaning is use*[145]. The meaning of a sentence is not determined by what the atomic parts of the sentence purport to refer to, but by the way the sentence (and its atomic parts) is used in ordinary discourse. Expressivism purports to explain the meaning of sentences in terms of the mental content expressed by speakers when they use the sentence in the light of sociolinguistic conventions. Thus, the expressivist framework reverses the order of semantic explanation as now mental content explains linguistic content (and not vice versa as in standard truth-conditional semantics). That is, thought explains language.

For example, Grass is green means 'Grass is green' because this is how the subject-predicate of the sentence is used to mean, according to the semantic and syntactic linguistic conventions that govern English language. That is, the semantic conventions that take grass to mean 'grass', green to mean 'green' and the syntactic conventions that regulate the role 'is' plays and the formation of grammatical sentences.

But as normative reference and truth-conditions do not necessarily show up in such an expressivist, use semantic picture, some use theorists like A.J.Ayer (1946), J.Austin (1962) and A.Gibbard (1990) have taken normative discourse to be non-truth-apt[146]. Reference and truth-conditions are simply explained away in their

---

[145] For some discussion of use theories see A.Avraamides (1999).
[146] J.Austin (1962) in particular takes all nondescriptive 'performative' discourse to be non-truth-apt and not just normative discourse.

picture of the semantics of normative discourse and this implies sliding to some form of normative relativism. Indeed, other use semantic theorists like H.Field (2009) have straightforwardly appealed to a relativistic expressivist conception of normative truth.

However, other use theorists like S.Blackburn (1998), A.Gibbard (2003) and R.Brandom (2000) have not been prepared to give up the precious notions of normative reference and truth-conditions and, more sanguinely, took them to be derivative of the use of normative sentences[147]. How to exactly gloss the notion of normative truth in such a use-theoretic framework is a very tricky question but their clear aim with this manoeuvre was to rescue the precious notion of normative truth.

And, arguably, such a manoeuvre is a wise one as there are deep philosophical reasons to think that the notion of normative truth is one that we should rescue, if our theory is to have any luck as a semantic theory of normative discourse. For a start, if we are to account for the validity of valid normative inferences, then we need some sort of expressivist-friendly account of truth as truth-preservation is necessary for validity[148]. But at any rate, all use theorists concur that reference should not taken to be the master concept of our semantic theory.

Expressivists usually rely on the seminal work of 'ordinary language philosophers' like J.Austin (1962) and, especially, P.Grice (1989). Gibbard (1990:85) relies, for example, on P.Grice's (1989) seminal work on ideational semantics. Along with J.Austin (1962), Grice's work on ideational semantics has been the starting point of contemporary work on ideational expressivism and has sparked a resurgence of interest to such a semantic approach[149]. But let me elaborate a bit on Grice and how his ideas are exploited by normative expressivists.

The central thing Grice (1989: 22-57) did was to famously explain how quite often the meaning of a sentence is not exhausted by what is merely being said. To show how this occurs he drew the distinction between *what is being said* and *what is being implicated*. According to Grice, what is being said is merely what is explicitly expressed by the atomic parts of the sentence, (that is, the words composing the sentence) in the light of the semantic and syntactic sociolinguistic conventions that

---

[147] I have in mind S.Blackburn's (1998) and A.Gibbard's (2003) Quasi-realism and R.Brandom's (2000) inferentialism.
[148] For a discussion of this point see M.Schroeder (2008: 161-3).
[149] See W.Davis (2010) for discussion.

govern the natural language in play.

For example, the descriptive 'snow is white' expresses the belief that snow is white according to the sociolinguistic conventions that govern English language. That is, the semantic conventions governing the meaning of words and the syntactic conventions governing the grammatical structuring of words into sentences[150]. Given that what is being said it is explicitly conveyed, this part of content is often called '*explicature*'[151].

But as Grice pointed out, quite often what is being said implicates things that go beyond what is merely being said, that is, the explicature. What is being implicated is what is being implied to an audience, according to the speaker's semantic intentions, over and above the sentential meaning of what is merely being said[152]. Given that such meanings are implicitly conveyed, Griced called this part of content '*implicature*'.

Grice distinguished between *conventional* and *nonconventional implicatures*. Conventional implicatures are the ones *non-cancellably* tied to what is being said, that is, the explicature. He gave this example: 'He is an Englishman, therefore, he is brave'. In this example, what is said also implicates that he is brave because he is an Englishman. This conventional implicature is expressed in all contexts that the sentence is uttered and therefore it is not cancellable. It is part and parcel of what is being said.

Instead, nonconventional implicatures are the *cancellable* meanings carried by explicatures, usually due to the speaker's semantic intentions, the conversational context and relevant background information. Other factors that might play a role in conveying nonconventional implicatures include intonation, facial expressions and overall body language.

Grice called such nonconventional implicatures *conversational* and distinguished

---

[150] 'Syntactic conventions' should be taken with a pinch of salt here because if N.Chomsky's (1980) work on the syntax of natural languages is to the right direction then syntactic conventions are not all that conventional. But not everyone is convinced from Chomsky's arguments. For arguments against 'syntactic universals' in natural languages see N.Evans and S.Levinson (2009).
[151] See W.Davis (2010).
[152] Audience should be understood in a technical rather colloquial sense in order to cover cases where the speaker speaks to himself. In such a case the speaker's audience is just himself. Grice (1989: 86-117) himself discusses how should one understand the notion of audience in cases where there is no *interpersonal* audience.

between *generalised* ones and *particularized* ones. Generalised conversational implicatures gradually tend to become conventionalised and when this happens they are rendered noncancellable[153]. Figures of speech like idioms, dead metaphors, proverbs etc. seem to be paradigmatic examples of such generalised conversational implicatures. For example, while the dead metaphor 'She has thrown oil on the flame' expresses what the sentence is literally saying, conversationally implicates that she has aggravated an already bad situation by means of fuelling more tension etc.

Particularized conversational implcatures are ones that arise often spontaneously out of the conversational context. Such implicatures might be entirely novel or of some limited use that has not yet allowed them to become generalized and conventionalized at the grander level of social circles or even society itself. The production of novel witticisms, subtle ironies, sarcastic and humorous expressions etc. might be good examples of such cases. For example:

John: 'Are we going to win the game?

Sally: 'Does the sun set in the east?'.

Sally's response here is an example of a particularized conversational implicature, somewhat cynically suggesting that their team does not really stand a chance and the possibility of winning is out of question. Such conversational implicatures are cancellable, as in a different context a speaker could utter the very same sentence without implicating that their team does not stand a chance and the possibility of winning is out of question. For example:

Sally the teacher: 'Does the sun set in the east?'.

Pupil: 'No, it sets in the west'.

In this case, Sally utters the same sentence but does not conversationally implicate anything germane to the chances of any team winning any game.

This short digression provides some of the basic ideas Grice worked out in his ideationalist framework. Grice, of course, didn't speak directly about normative sentences as such which is the object of our interest here[154]. The closest he comes in saying something relevant about how his approach could apply to normative

---

[153] When conversational implicatures became generalised and conventionalized and therefore noncancellable they are, in essence, rendered conventional implicatures.

[154] Austin (1962) only mentions in passing legal, epistemic and ethical discourses and in his (1961) elaborates some of his ideas to epistemic discourse.

discourse is when he (1989: 213-224) briefly makes the distinction between descriptive (or 'informative') sentences and imperative sentences, as normative sentences resemble imperatives in interesting ways[155]. But this has not really hindered Gibbard from borrowing and applying Gricean ideas on the semantics of the normative, rationality discourse himself.

As we will shortly see in the next section 6.3, Gibbard is following Grice's work on ideational semantics when he proposes that when we express normative sentences we are not merely saying something but, importantly, we are also implicating things. For Gibbard, the class of normative sentences are one class of sentences where what is being said does not exhaust their meaning because such sentences also implicate things. Indeed, for Gibbard what we *really* mean when we express normative sentences is not what is being said but what is being implicated.

Gibbard makes this claim because he thinks that the *surface-level* referential pretensions of normative language are deceptive and misguiding. They are deceptive because, as we saw ourselves in Chs.2-5, is far from clear what the epistemic justification property (or other normative properties) could be or how we could track such a property and this inspires scepticism about the prospects of referential semantics. For Gibbard, the real meaning of these sentences lies at the *bottom-level* of what is being implicated and what is implicated, according to ordinary rationality discourse, is a non-cognitive mental state, namely, an attitude of norm-endorsement[156]. In J.Austin's (1962: 4) terms such sentences are 'masquerades', they are sentences in referential syntactic disguise which, nonetheless, are not semantically referential.

With some of the basic ideas of the semantic programme of expressivism (and its Gricean underpinnings introduced), let us now turn to A.Gibbard's norm-expressivism about moral rationality.

6.3 A.Gibbard's Moral Norm-Expressivism

In his influential book *Wise Choices, Apt Feelings,* A.Gibbard (1990:4-5) sets out searching for an understanding of 'what 'rational' means' and how an understanding

---

[155] For discussion of how normative sentences resemble imperative sentences see R.M.Hare (1952: 1-31) and M.Schroeder (2008: 10-12).
[156] The surface-level/bottom-level distinction here parallels L.Wittgenstein's (1953: 168e) 'surface grammar/ depth grammar' distinction.

of the meaning of 'rational' could subsequently elucidate the species of moral rationality. That is, what makes our moral choices to 'make sense' or 'to be wise'. He (1990:6-12) patently renounces a *referential semantic strategy* in favour of an ideational *use semantic strategy* as an approach to the meaning of rationality assertions and attributions. That is, he rebuffs referentialist efforts to either reduce rationality to any naturalistic property or to assert that there is an irreducible nonnatural rationality property.

Instead, he suggests that we should start with how rationality assertions and attributions are actually *being used* in ordinary linguistic practice. For Gibbard, what ordinary linguistic practice reveals is that with normative sentences we express noncognitive mental states, namely, attitudes of norm-endorsement. That is, attitudes of endorsement for the moral norms that license the action in question as rational. He calls this approach to the semantics of rationality discourse *norm-expressivism*. Let me elaborate on how exactly he arrives at norm-expressivism.

Gibbard starts with a subtle dialectic manoeuvre. Gibbard's dialectic manoeuvre is to crucially *reverse* the order of semantic explanation. Referentialists-reductionists start with alleged analyses of the nature of normative properties like rationality, goodness etc. and think that endorsement will simply follow from the correct analysis. Once we have the correct analysis of such normative notions, they presume, endorsement will nicely fall in place. But due to the familiar 'Moorean/Humean lesson' this does not seem to be forthcoming.

In the face of 'the Moorean/Humean lesson' and the pessimism it inspires, Gibbard reverses the order of semantic explanation[157]. He starts with the use of these sentences and the element of endorsement and sets out to explain norms of (moral) rationality as something following from this use. That is, he sets out from our attitude of endorsement to what we take to be rational and how this endorsement seems to be licensed by certain norms. Norms that constrain what can count as rational and what cannot. These norms, however, are neither somehow meant to be 'out there' nor they are meant to be mind-independent and objective in the traditional realist fashion. They are norms following from the use of justification assertions and attributions by epistemic agents.

---

[157] Gibbard, as a philosopher that adheres to methological naturalism, he is not willing to go nonnaturalist by any means.

With Gibbard's reversal of the order of explanation and the invocation of an ideational 'use semantic strategy', the notion of norm-endorsement comes to lie at the heart of the proposal. Gibbard claims that the use of such normative sentences indicates the importance of the element of norm-endorsement and that this should be the *cornerstone* for a theory of the semantics of rationality assertions and attributions. As he puts (1990:6) it:

> '...start with the use of the term. Fix on the dictum 'To call a thing rational is to endorse it', and search for a sense of 'endorse' for which the dictum holds true'. He then adds (1990:7) that 'to call something rational is to express one's acceptance of norms that permit it'.

According to Gibbard, the meaning of 'rational' should be understood in terms of the noncognitive state of mind it expresses, namely, the nonreferential attitude of acceptance (or endorsement) to norms that permit the action, event or state of affairs in question.

Gibbard has some very good reasons for making the notion of norm-endorsement the cornerstone of his expressivist theoretical construction. For, on the one hand, as we have seen in section 2.3, referentialist approaches seem to exactly *miss* the element of endorsement[158]. As you might recall, 'the Moorean/Humean lesson' suggested that it is always possible for a clear-headed agent, without any semantic confusion, to question whether any normative property N (e.g. rationality) is reducible to a natural property N* (e.g. full information, desire satisfaction) and therefore resist endorsement of the belief in question. I can see, he might retort, the full information about, let us say, the Armenian genocide facts but I can't see why full information makes them rational.

As Gibbard (1990:10) puts it: 'On that diagnosis, what descriptivistic analyses miss is a general element of endorsement- an element an expressivistic analysis can capture'. And an expressivistic analysis can capture this element of endorsement because it reverses the order of explanation and starts with what agents endorse when they use rationality assertions and attributions. What they endorse when they state that 'φ-ing is rational' or 'S's φ-ing is rational'. Thus, the endorsement element in the norm-expressivist story allows us to account for what we may call 'the

---

[158] See also Gibbard (1990:10).

endorsement intuition', namely, the intuition that when we use rationality assertions and attributions we seem to express some sort of general endorsement.

On the other hand, Gibbard is attaching to endorsement the 'norm-' prefix because a sense of 'endorse' should have something to do with norms. Intuitively, norms seem to constrain what can count as rational (moral or epistemic) and what cannot and Gibbard is sensitive to this fact. Wisely, although wants to dispense with 'robust' norms or facts (realistically construed) about rationality, he does not want to relinquish the normative element inherent in rationality assertions and attributions.

As he (1990:8) says :

'In many ways, normative judgements mimic factual judgements, and indeed factual judgements themselves rest on norms...Normative discussion is much like factual discussion, I shall be claiming, and just as indispensable'.

Thus, the 'norm-' prefix allows us to account for what we may call 'the norm-constraining intuition', namely, the intuition that our rationality assertions and attributions are being constrained by norms.

This concise sketch clarifies a bit the ideational semantic underpinnings of Gibbard's norm-expressivism and delivers its contours. But a last point is due before we conclude this section. It is important not to fail to appreciate what motivates Gibbard's overall norm-expressivist project. His motivation for an ideational use semantic strategy and a norm-expressivist account of the semantics of moral rationality springs from his strong commitment to *methodological naturalism*. Gibbard, as a philosopher with methodologically naturalistic orientation, intends to firmly situate the genus of rationality (and the species of moral rationality) in nature. He wants to understand in thoroughly naturalistic terms how we, human animals, have possibly come to be the animals with the marvellous practical and cognitive capacities we have.

To that effect, he also speculates about the naturalistic origins of our capacity for norm-endorsement. His intriguing speculations suggest that the human capacity to 'accept or endorse norms' has been the product of biological adaptation through Darwinian natural selection[159]. This capacity has been adapted through natural

---

[159] Gibbard's (1990: 29-30, 256) invocation of evolutionary theory is subtle. He is well aware of the complexities surrounding evolutionary theory and his approach cannot be seen, I think, as a form of crude adaptationism.

selection for serving a certain evolutionary advantageous functional purpose.

The functional purpose of our capacity for judgemental endorsement of norms is to coordinate interpersonal interaction in a way that would enhance our reproductive success and survival chances. It is to coordinate social life and cooperation often with mutual benefit for the persons involved. In line with this thoroughly naturalistic picture, Gibbard also speculates about the evolutionary origins of our 'emotional propensities', often involved in expressing our attitudes of norm-endorsement. According to Gibbard, they are largely the consequence of our evolutionary history and had been adapted to function as social lubricants that regulate coordination and guide social co-operation.

Now, my intention in this chapter 6 is to borrow Gibbard's (1990) norm-expressivist framework for moral discourse and with some twists apply it to epistemic discourse and, in particular, to epistemic justification which is the specific object of my inquiry. I will use it as an approach to 'the epistemic justification puzzle' and explore what explanatory fruit such an approach could reap.

It should be made clear, however, that as far as the question of the origins of our capacity for norm-endorsement is concerned, I won't mimic Gibbard's evolutionary speculations and choose to demur. It would have been deeply interesting to parallel Gibbard's speculations about the evolutionary origins of our capacity for moral norm-endorsement and speculate about the evolutionary origins of our capacity of epistemic norm-endorsement, but this parallel would lead us too far from what the scope of this work allows.

The idea, though, is bold and challenging and not obviously misguided or incoherent. Indeed, there are philosophers who speculate along those evolutionary lines like R.Millikan (2003), D.Papineau (2003) and, also, Gibbard's (2003) own later work on epistemic plan-reliance expressivism fits nicely in such a context. Unfortunately, I will have to forgo here the chance for speculative evolutionary ruminations on our capacity for norm-endorsement and will remain neutral about its provenance.

With Gibbard's (1990) norm-expressivist semantics about rationality sketched, in the next section we apply a version of norm-expressivist semantics to epistemic justification discourse.

6.4 From Moral Norm-Expressivism to Epistemic Norm-Expressivism

In this section I apply a version of norm-expressivist semantics to epistemic justification discourse. This application has already been executed by H.Field (2000, 2009), though, Field executes the application towards what he (2009) calls 'an expressivist relativist' direction that, for reasons I shall explain, I resist here. This also marks a departure from Gibbard's own gloss of the theory because Gibbard (1990), like other early expressivists, took normative discourses to be non-truth-apt and this seems to imply sliding to a form of relativism that, as I said, I want to resist.

In the footsteps of Gibbard, the norm-expressivist proposal starts from the claim that a *referential semantic strategy* as an approach to the meaning of justification assertions and attributions is misguided. As a result, the 'referential semantics assumption' (P2) is false and this is the way out of the puzzle. When we use justification assertions and attributions we don't purport to refer to a certain epistemic justification property (natural or nonnatural).

Despite the surface-level referential pretensions of justification assertions and attributions, the semantic function of such normative sentences is not referential at all. What is being said when we express normative sentences should not be allowed to deceive and misguide us. Instead, the semantic function of normative sentences is expressive of a noncognitive and nonreferential mental state. The real meaning of such sentences lies at the bottom-level of what is being implicated. As a result, we should neither strive to reduce the notion of epistemic justification to necessary and sufficient conditions nor postulate a sui generis property of epistemic justification. We should not postulate the existence of an epistemic justification property at all.

Rather, it suggests that we should employ an ideational 'use semantic strategy' for understanding the meaning of justification assertions and attributions. This is the paramount dialectical manoeuvre for building a theory of semantics for such sentences. We should start with how epistemic justification assertions and attributions are actually *being used* in ordinary discourse. The idea is that the actual use of such sentences in ordinary discourse will reveal to us what determines the meaning of these sentences.

Following Gibbard (1990), the core ingredient the use of normative sentences seems to reveal, and a theory of the semantics of such sentences should better take

seriously, is the feature of *'endorsement'*[160]. That is, when epistemic agents state that 'p is (epistemically) justified' or that 'S justifiedly believes that p' they seem to express some sort of general endorsement (or acceptance/approval) for the belief in question as being (epistemically) good. The agent seems to approve the belief as if it is something (epistemically) good or valuable.

All things considered, epistemic agents feel that holding a justified belief is something that is (epistemically) good and praiseworthy, something they can take credit for. This is because epistemic justification seems to be necessary for knowledge (broadly construed) and, generally speaking, we think that knowledge is something desirable because it is (both instrumentally and finally) valuable[161].

But with justification assertions and attributions epistemic agents don't just express endorsement for the belief *simpliciter*. They seem to express '*norm-endorsement*'. The 'normative' character of endorsement should be understood in the sense that epistemic agents implicitly endorse certain epistemic norms that license the belief in question. For, clearly, by the agent's lights not just any belief can be endorsed as justified. By the agent's lights, there should be some epistemic norms constraining what can be taken as justified and what not.

These comments make clear that the feature of norm-endorsement is one of the basic desiderata that a theory of the semantics of justification assertions and attributions must accommodate. This is exactly what we do. Following Gibbard's (1990) norm-expressivist semantics, the core element of norm-endorsement lies at the very heart of this proposal and allows us to vindicate two central intuitions about epistemic discourse, namely, 'the endorsement intuition' and 'the norm-constraining intuition'. That is, it allows us on the one hand, to easily capture the endorsement element that referentialist approaches seem to exactly miss, as we have already discussed above and seen in Chs.2-4, and on the other hand, to explain how norms constrain epistemic discourse (though not in the traditional realist sense).

Taking the norm-endorsement ingredient to be the cornerstone of our approach, this is then a first rough and general statement of the norm-expressivist proposal. The

---

[160] See A.Gibbard (1990:6-10).

[161] On the problems surrounding the value of knowledge see D.Pritchard (2007, 2010) and K.Kappel (2010). In my 'Evolutionary Ruminations on 'the Value of Knowledge Intuition'' (forthcoming), I also provide some tentative evolutionary psychological ruminations on why we are so effortlessly inclined to find knowledge valuable.

semantics of epistemic justification assertions and attributions are expressivist and the state of mind expressed noncognitive .i.e. attitudes of norm-endorsement. That is, attitudes of endorsement (or approval) of the epistemic norms that license the belief in question.

When we make justification assertions and attributions we don't purport to refer to a certain natural or nonnatural 'epistemic justification' property. Rather we express nonreferential mental states, namely, noncognitive attitudes of endorsement for the norms that license the belief as justified. By analogy, when we make unjustifiedness claims and attributions we again express attitudes of norm-endorsement. That is, attitudes of endorsement for the epistemic norms ruling out the belief from being justified.

This rough norm-expressivist picture, obviously, implies the rejection of the key theoretical commitments of the referentialist positions we have examined in Chs.3-5: 'the referential semantics assumption', 'the realist-ontological assertion' and 'the existential assertion'. As we have seen, the assumption that justification assertions and attributions purport to refer to an epistemic justification property gives rise to the further ontological assumption that there is such a property. Realists think that there must be such a property because justified beliefs, inferences etc. should be justified in virtue of a property of epistemic justification. This is what we have called 'the realist-ontological assertion'.

Yet, as I have taken pains to argue, it seems very tricky to see what this property (natural or nonnatural) could ever be. For one thing, we have noted that the 'in virtue of' seems to be causal and it is not clear how such a normative property could have independent causal powers. The rejection of 'the realist-ontological assertion' seems to also imply the rejection of 'the existence assertion', namely, that there are at least some justified beliefs, theories etc. This implication might seem unpalatable because no one -perhaps with the exception of some sceptic -not just about knowledge but even about epistemic justification- would welcome such a conclusion. Almost everyone seems to think that there are at least some justified beliefs, theories etc.. Besides, even some sceptics might concede that scepticism about knowledge is legitimate but not legitimate about justification[162].

---

[162] As it is sometimes noted, there is less pressure to take scepticism about justification seriously than scepticism about knowledge. See R.Wedgwood (2009).

But from the rejection of 'the referential semantics assumption' and 'the realist-ontological assertion' the norm-expressivist insists that we should not hasten to infer that there are no justified beliefs. At least we should not hasten to infer that there are no justified beliefs in some interesting sense of 'justified'. Admittedly, there are no justified beliefs in the traditional realist sense that demands a property of epistemic justification in virtue of which justified beliefs are being justified. This understanding of 'the existential assertion' we may call 'the *realist* existential assertion' and it is the one the realist has been assuming all along. But the norm-expressivist can settle for an ontologically light assertion: 'the *minimalist* existential assertion'. Let me explain.

As we have seen, the realist thinks that there are at least some justified beliefs (i.e. this is the initial '(realist) existential assertion') and that in order to account for these we need to postulate a property of epistemic justification in virtue of which justified beliefs are justified. Norm-expressivism contests this claim and suggests that we can have justified beliefs even *without* any epistemic justification property. We can have justified belief in some interesting sense of 'justified', though, not in the full-blooded ontologically burdened sense of the realist.

Our justified beliefs etc. will be justified due to the norm-governed use of justification assertions and attributions, not due to the existence of an elusive property of epistemic justification. For the norm-expressivist, the epistemic practice of using justification assertions and attributions is something following from plain norm-governed linguistic activity without any extra ontological and meta-epistemological baggage.

Here is an example of such a minimally justified belief. Suppose I say: 'I justifiably believe that Hannibal was a great strategist'. According to our norm-expressivist story what I express is a noncognitive attitude of endorsement of the epistemic norms that license this belief as justified. Many different sorts of epistemic norms (foundational, coherentist, virtue-based etc.) are eligible to license this belief as justified but let us say for the sake of exposition that my epistemic norms are foundational. In that case, I express an attitude of endorsement for the foundationalist epistemic norms that license as justified my belief that Hannibal was a great strategist.

In the light of my foundationalist norms, I might find this belief 'self-presenting' due to other background beliefs about what it takes to be a good strategist and Hannibal's military exploitations. I might believe, for example, that any strategist that has mastered and employed the pincer movement must have been a good one and also believe that Hannibal masterfully used that stratagem at the battle of Cannae. Therefore, in the light of these background beliefs I find the belief that Hannibal was a great strategist 'self-presenting' and thus justified.

Of course, there might be minimally justified beliefs where the agent's epistemic norms are simply fanciful and whimsical from an epistemic point of view. That is, from the point of view of being truth-conducive, as some agent's epistemic norms might not be truth-conducive in any sense. Suppose for example that an agent makes justification assertions and attributions on the basis of the epistemically irrational norm of blind hunches. He claims: 'I justifiably believe that Gaius Terentius Varro was a great strategist'.

According to our norm-expressivist story, he expresses an attitude of endorsement for the epistemic norms that license this belief, namely, the epistemic norm of blind hunches that licenses this belief as justified. He might think, for instance, that Gaius Terentius Varro is a cool-sounding name and therefore it is intuitive that he must have been a good strategist by all means. This is another case of a minimally justified belief licensed, though, by epistemic norms that are not just as truth-conducive as in the former example.

Having minimally justified beliefs, though, that permit at least some of these beliefs to be not truth-conducive in any interesting sense raises the legitimate question of how revisionist is this norm-expressivist conception of epistemic justification we have arrived at. That is, how much it departs from the ordinary notion of epistemic justification as we have discussed it in section 2.2.

For as we canvassed in section 2.2, a theory of epistemic justification should satisfy 'the truth-conducivity intuition', namely, the intuition that epistemic justification is truth-conducive as it makes beliefs likely to track the truth. Yet norm-expressivism seems to allow that many norm-justified beliefs are not truth-conducive as they are being licensed by merely fanciful epistemic norms. This, at first instance, suggests that minimally justified beliefs are not really justified as justified beliefs

should be truth-conducive and at least many minimally justified beliefs are not to be considered truth-conducive.

This is a serious challenge for norm-expressivism as it threatens to provide a theory of epistemic justification too revisionary for what our ordinary notion of epistemic justification allows. One direction of response here is to bite the bullet like H.Field (2009) and appeal to a norm-relative conception of truth and be a happy 'expressivist relativist', as he says. But I think this response is inadequate for two reasons.

First, it departs from what we seem to mean when we make justification assertions and attributions, as ordinary normative discourse is pervasively realist-seeming. When we make justification assertions and attributions we seem to do so from a realist standpoint, namely, we seem to imply that there is a fact about what is justified and what is not. Relativism, therefore, seems to be too revisionary for one like me that takes realist-seeming intuitions seriously and wants 'to save the epistemic phenomena'.

The second point follows, really, from the first. If we want to capture our realist intuitions about justification discourse but seem to meet difficulties, as we have seen in chs.2-5, then maybe we can capture these intuitions from the inside out with the antirealist standpoint of expressivism. Indeed, this is what many proponents of expressivism think is its great promise, namely, to account for our realist standpoint in normative discourse in an antirealist fashion[163]. Gibbard (2003), for example, goes as far as to mimic Kant and talk of expressivism as effecting 'a Copernican revolution' to philosophy because he thinks expressivism can explain the realist seemings of normative discourse in an antirealist framework[164].

Based on these considerations, a better response, I think, is to build on the broadly social, interpersonal realm. We could suggest that minimally justified beliefs that are based on irrational, non-truth-conducive norms can be criticized from an interpersonal, social point of view. Agents employing irrational norms like premonitions, hunches, lucky guesses etc. are to be castigated from the social point of view because these norms can't function as a means to the social good of truth. Instead, coherentist, foundational, virtue-based etc. norms that, although do not

[163] See S.Blackburn (1993, 1998), M.Timmons (1999) and A.Gibbard (2003).
[164] Austin (1962: 3-4) is equally enthusiastic and talks of '...producing a revolution in philosophy'.

reduce epistemic justification, can often function as a means to truth are to be praised and encouraged. There is more to be said about this aspect of epistemic justification discourse but this is sufficient for present purposes. I will return to the issue of epistemic improvement in the next chapter 7 when I will introduce my proposal.

A last point remains before we pass to the next section. As I said in section 6.1, Gibbard (1990) (like Ayer and Austin) took normative discourse to be non-truth-apt. Normative sentences are not to be assigned truth-conditions and be evaluated in terms of truth and falsity. Yet, as I have already hinted in section 6.1, there are deep philosophical reasons to avert us from such a move. I need not expand on these reasons here but a key one is that if we are to hope to account for the logical properties of normative sentences as found embedded in unasserted truth-functional contexts (conditionals, disjunctions etc.), then we require some sort of expressivist-friendly account of truth for normative discourse.

Gibbard (2003) himself in later work recants from taking normative discourse to be non-truth-apt and follows Blackburn's (1993, 1998) appeal to minimalism about truth and his so-called program of 'quasi-realism'. The 'quasi-realism' idea is that we can have normative realism in an antirealist fashion, that is, without the ontological and epistemological package of the realist that seems to have the familiar dire consequences. Hence, the 'quasi-' prefix on realism. In this way, Blackburn and Gibbard suggest we can take normative discourse to be truth-apt and therefore evaluable in terms of truth and falsity, even if as antirealists-expressivists we don't postulate normative properties.

I won't expand much on how we could rescue normative truth in such a use-theoretic framework but I can hint that one approach that seems prima facie appealing to me comes from R.Brandom (2000). Brandom has argued that we can ground normative truth and objectivity on the social level, at the level of social articulation of reasons for or against a normative claim. As I understand him, his idea is that normative truth is something that is being born at the social level where the transaction of reasons takes place. As such, normative truth is the outcome of a social and historical process and takes time and argument to crystallize and show up. Similar ideas in some respect are being suggested by Blackburn (1998: 200-212), though, Blackburn is inclined to trace the idea of grounding normative truth and

objectivity at the social level to A.Smith and D.Hume while Brandom back to Hegel.

For example, 'slavery is wrong' or 'heliocentricism is justified' are normative claims that came to be seen as undoubtedly true because of a social and historical process that took time and argument. Reasons for and against were given and in time, as passions and prejudices subsided, it became clear that the reasons for had by far the upper hand. A convinced realist might object here that this is no better than social-cultural relativism but I will leave the issue at this point. All I mean to make explicit here is that I don't follow Gibbard's jettison of normative truth and I do acknowledge the need for an expressivist-friendly account of truth. How this is exactly to be achieved it remains to be explored and hopefully to be shown.

Admittedly, it is not immediately evident that such an antirealist conception of normative truth will be in the end defensible nor that it won't be as there is fierce disagreement surrounding 'truth debates' but, in any event, it is good to make clear what commitments our approach needs to take aboard. Whether it can eventually discharge these commitments is an independent question[165].

With the emulation step now completed, in the next section 6.5 I step forward to examine how our 'toy' norm-expressivist theory can easily evade the referentialists' discontents and, therefore, motivate the possibility of an expressivist approach to the epistemic justification discourse.

6.5 Motivating Expressivism: How Expressivism Can Evade Referentialist Problems

As we already said in section 6.1, the expressivist approach to 'the epistemic justification puzzle' seems to easily evade the referentialists' problems we presented in the critical Chs.2 -5 and this motivates the possibility of an expressivist approach to epistemic justification discourse. I will here use our version of norm-expressivism as a 'toy' theory in order to explain how the expressivist theory can evade these problems. As I said, this is just for the sake of expounding how expressivist theories can evade the referentialists' discontents.

First, contra the analytic naturalistic reductionist, norm-expressivism seems to explain the existence of Moorean/Humean 'open feel' semantic intuitions that thwart

---

[165] Some people argue that it can't. See Kvanvig (2003), Cuneo (2007) Lynch (2008). But for expressivist responses see K.Kappel (ms) and M.Crisman and A.Carter (ms).

attempts to reduce the notion of epistemic justification to a natural property. The norm-expressivist suggests that the existence of our Moorean/Humean intuitions can be *best* explained as due to the inexistence of such a property. There is no natural property capable of reducing epistemic justification and, hence, the notion should be considered to be unanalyzable. For the expressivist, justification assertions and attributions do not purport to refer to such a property at all. They express attitudes of norm-endorsement and this follows from the plain norm-governed use of epistemic concepts in ordinary linguistic activity.

Second, contra the synthetic naturalistic reductionist and the epistemic kinds realist (under a naturalistic reading), it seems to explain 'the epistemic twin earth' semantic intuitions that twhart attempts to either reduce epistemic justification to a natural kind or postulate it as a natural 'epistemic kind'. For the norm-expressivist, our persistent 'open feel' semantic intuitions can be *best* explained as due to the inexistence of such a natural kind or a natural 'epistemic kind' property of epistemic justification. This is why our 'open feel' semantic intuitions persist unabated independently of whether we talk of reduction by semantic analysis, reduction by appeal to a natural kind or reduction by an elusive natural 'epistemic kind'.

Third, contra the nonnaturalist, the norm-expressivist approach seems to avoid the triumvirate of its discontents. On the one hand, it seems to provide a naturalistic account of epistemic justification and thus eschews 'the epistemic access' and 'queerness' problems that beset nonnaturalist approaches. For the expressivist, justification assertions and attributions do not purport to refer to a 'queer' sui generis property of epistemic justification. They assert that there is no such 'queer' property and take justification assertions and attributions to express attitudes of norm-endorsement.

Equally, for the expressivist there is no special cognitive faculty of rational intuition tracking the property of being justified as instantiated by beliefs, theories etc.. All we have as human beings are conceptual powers, bestowed to us through language acquisition in the social setting of a community and these conceptual powers enable us to express our attitudes by means of using words and sentences. In the case of justification assertions and attributions sentences, we express our attitudes of norm-endorsement, that is, our attitudes of endorsement for norms that license the

belief in question as justified.

In addition, contra nonnaturalist approaches, norm-expressivism seems to explain the 'epistemic supervenience desideratum'. As it is often pointed out, and as we have seen, any theory of the semantics of epistemic discourse must explain how epistemic properties (e.g. being epistemically justified or being epistemically rational) supervene on non-epistemic, natural properties (e.g. coherence, self-presentation etc.)[166]. That is, how two distinct but naturalistically identical situations (that is, sharing the very same natural properties) cannot realize different epistemic properties. In slogan form, there can be no epistemic difference without a natural difference.

To contest supervenience is to accept that there could be a possible scenario where two distinct but naturalistically identical situations realize different epistemic properties; something that appears absurd. For example, in two distinct but naturalistically identical situations, if in the first situation natural facts (fingerprints, eyewitnesses' testimony, motives etc.) justify the belief that there has been murder then in the other situation it would have been absurd to think that the same natural facts justify, instead, the belief that there has been suicide.

One way to explain epistemic supervenience is in the ontological terms just formulated above, despite the fact that for the expressivist there are no epistemic properties. The expressivist could still 'speak with the vulgar' as if there are epistemic properties and suggest that two distinct but naturalistically identical situations should not differ in terms of realization of epistemic properties. For the expressivist, if the situations are naturalistically identical then the epistemic agent should express the same epistemic evaluation. Otherwise, he would breach the supervenience desideratum and be charged with irrationality[167].

A better way for the expressivist to explain epistemic supervenience is to follow R.Hare (1952) and A.Gibbard (2003) and spell out the notion of supervenience in terms of concepts, not properties[168]. This approach would allow us to sidestep the gratuitous invocation of the Humean notion of 'speaking with the vulgar' about an epistemic justification property.

---

[166] See J.Kim(2008); R.Feldman (2001); H.Vahid (2005).

[167] See M.Ridge (2003) for taking this line of an expressivist explanation of moral supervenience.

[168] This is what J.Klagge (1988) has called 'ascriptive' instead of 'ontological supervenience'.

According to this concepts-based conception of epistemic supervenience, supervenience constrains how we should *use* epistemic concepts and not how epistemic properties are realized, as for the expressivist there are no such properties. If two distinct situations are found to be naturalistically identical then, our epistemic evaluation should be the same in both situations. Otherwise, our use and application of epistemic concepts will threaten to verge on the irrational and arbitrary for the reasons just explained above.

In conclusion, expressivism easily outflanks the referentialists' problems we identified in Chs.2-5 and this shows that expressivism is an approach that we should take seriously. Expressivism has certain important explanatory advantages compared to theoretical competitors that entitle one to pursue it in good hope.

6.6 Conclusion and Summary of the Argument

This chapter 6 had a twofold dialectical purpose: to introduce the basic ideas of the semantic program of expressivism and motivate the possibility of an expressivist approach to the epistemic justification discourse by means of showing how a 'toy' version of norm-expressivism could easily evade the problems that have made the life of referentialist approaches so hard in the chs.2-5. In section 6.2, I introduced the basic ideas of the semantic programme of expressivism. As I said, unlike current orthodoxy in semantics which is referential and truth-conditional, expressivists explain the meaning of sentences in terms of the mental content expressed by the speaker in the light of sociolinguistic conventions. Mental content thus is semantically primary and accounts for linguistic content too. Thought explains language, if you want, and not vice versa.

In section 6.3, I introduced the key ideas of A.Gibbard's norm-expressivism for moral rationality discourse and in section 6.4 applied it to epistemic justification discourse. Unlike Gibbard (1990) himself who said that rationality discourses are not truth-apt and H.Field (2009) who has taken norm-expressivism to be a norm-relative theory I have suggested that expressivism could legitimately aspire to cling to our realist intuitions, though, in an antirealist fashion. To that effect, I have hinted to R.Brandom's (2000) grounding of objectivity to the social realm of the game of giving and asking reasons for our claims.

Finally, in section 6.5 I motivated the possibility of expressivism by explaining

how our toy expressivist theory can evade the problems of the referentialist theories we have witnessed in chapters 2-5. As I have explained our toy norm-expressivism can account for our Moorean/Humean 'open feel' semantic intuitions inherent in 'the Moorean/Humean lesson' and 'the epistemic twin earth lesson', evade 'the epistemic access' and 'queerness' problems for the nonnaturalist and, in addition, account for 'the epistemic supervenience desideratum' that nonnaturalists fail to accomodate.

The ability of expressivism to evade the problems of the referentialists so easily motivates the conceptual exploration of an expressivist approach to epistemic justification discourse and more generally the normative domain. For if expressivism can reap this explanatory fruit so easily that other theories struggle with, then expressivism is a potentially promising approach that needs to be explored to some depth.

In the next chapter 7 I do exactly this. I explore how an original version of norm-expressivism about epistemic justification discourse should go. The original version of norm-expressivism I arrive at I call 'habits-endorsement expressivism'. I motivate this new version of norm-expressivism by means of arguing that it can evade two identified problems for a plan-reliance expressivist approach to justification discourse while also embrace the virtues of both norm-expressivism and plan-reliance expressivism. Let us now turn to this more constructive task of theory-building.

Chapter 7 The Possibility of Habits-Endorsement Expressivism

7.1 Introduction

As we have argued in Ch.6, our 'toy' version of norm-expressivism easily evades the referentialist's puzzling problems and this *motivates* the possibility of a broadly expressivist approach to 'the epistemic justification puzzle'. At least, it motivates a conceptual exploration of an expressivist approach in good hope that this will prove both interesting and explanatory fruitful. Yet, even if we grant that a broadly expressivist approach to the puzzle should be explored in good hope, it remains unclear what sort of expressivist story should that be, as a number of competing expressivist approaches to epistemic discourse can be found in the recent literature[169].

From a methodological point of view, no doubt, the best expressivist theory should be the one that is the most explanatorily fruitful in -by fiat of Occam's razor- the most simple and economic way possible. So, what we need is an expressivist theory that not only evades the referentialists' problems but also explains *more* key intuitions about epistemic justification discourse than its competitors in the most simple and economic way possible. That is, it explains intuitions that seem to be interwoven in the way we make justification assertions and attributions.

This is what I intend to explore in the current chapter. I intend to explore an expressivist approach to epistemic justification discourse that not only evades the referentialists' problems but also promises to account for key intuitions about epistemic discourse that at least some other prominent expressivist theories do not. This hopeful new expressivist theory I call 'habits-endorsement expressivism' and it is in reality an enhanced version of early Gibbard's (1990) norm-expressivism.

But as I can't pursue at length here a comparative analysis of my proposal with all these expressivist theories, I will provide some motivation for my proposal by means of showing how it can evade some identified problems of A.Gibbard's (2003, 2008) more recent plan-reliance expressivism. Or, to be accurate, A.Gibbard's plan-reliance

---

[169] See A.Gibbard (2003; 2008), H.Field (2000; 2009), M.Chrisman (2007), M.Ridge (2007b), K.Kappel (2010; ms).

expressivism as extended to cover epistemic justification discourse. As we have seen in Ch.6, early Gibbard (1990) had advocated norm-expressivism about moral rationality discourse that has been relatively influential. Yet, in later work (2003, 2008) he went on to develop a different, plan-reliance form of expressivism for both moral and knowledge discourse and there are good reasons to suspect that plan-reliance expressivism, even if it works for knowledge discourse, it cannot be transposed to epistemic justification discourse with impunity.

It cannot be transposed to epistemic justification discourse with impunity because, even on the (controversial) assumption that we have so far strived to motivate, namely, that epistemic justification discourse is to have an expressivist semantic treatment, this theory faces problems that undercut its prima facie promise. Instead, as I will suggest, the old norm-expressivist version of expressivism with some helpful tweaks can do much better in regard to these problems. Let me explain.

On the one hand, plan-reliance expressivism was introduced by Gibbard (2003, 2008) as a theory of the semantics of knowledge discourse and not the justification discourse and the problem is that, even if we grant the (controversial) assumption that it does work for knowledge discourse, this approach runs into direct counterexamples that impede extension to epistemic justification discourse[170].

The problem is that Gibbard's plan-reliance expressivism suggests that knowledge assertions and attributions express plans of reliance, namely, plans to rely on the belief in question and this seems clearly *mistaken* when applied to epistemic justification discourse. It seems clearly mistaken because we often attribute justified belief but do not plan to rely on this belief because we take it to be, nonetheless, false. Such cases constitute counterexamples to plan-reliance expressivism and impede extension of application from knowledge to justification discourse. Thus, plan-reliance expressivism as an approach to the semantics of overall epistemic discourse remains incomplete, as it can't extend to cover an important segment: the epistemic justification discourse.

On the other hand, Gibbard's plan-reliance expressivism when applied to epistemic justification discourse also does not seem to account in any direct sense for

---

[170] See M.Ridge (2007a) for criticism of Gibbard's (2003) plan-reliance expressivism for knowledge discourse.

a key intuition of epistemic discourse: 'the nonvoluntarism intuition'[171]. The nonvoluntarism intuition draws support from everyday doxastic experience to suggest that (justified) belief-fixation is not subject to direct voluntary control. We can't choose at will what (justified) beliefs to form as (justified) belief-fixation seems to be nonvoluntary. But plan-reliance expressivism is build on the notion of normative plans of reliance, which has strong voluntary connotations, and this seems to suggest that we can choose our epistemic plans, namely, plans about what to believe and rely on. Yet, this is something that the voluntarism intuition denies and plan-reliance expressivism seems to defy the intuition without an explanation.

However, in spite of these problems things are not all that bleak for expressivists. The good news is that philosophers sympathetic to expressivist approaches to epistemic discourse should see no reason to despair in the face of these problems because, arguably, we could have the best of both worlds. That is, we could have an expressivist theory that both keeps the attractions of early and late Gibbardian stories while at the same time avoids the problems of plan-reliance expressivism just mentioned above. Or at least this is what I will argue for here.

I open the discussion with Gibbard's (2003, 2008) plan-reliance expressivism for knowledge discourse and then outline how it would look like if it were extended to cover the epistemic justification discourse. I then explain how plan-reliance expressivism runs into direct counterexamples and remains incomplete as an overall theory of the semantics of epistemic discourse whilst it also does not account for the nonvoluntarism intuition.

The final step will be to constructively synthesize insights in an original version of norm-expressivism that, on the one hand, captures what is attractive in norm-expressivism and plan-reliance expressivism and, on the other hand, outflanks the problems plan-reliance expressivism meets when applied to epistemic justification discourse. At the core of the proposal will be laid the notion of *epistemic habits*, as exploiting this versatile and pregnant notion will allow us to both outflank the counterexamples plan-reliance expressivism encounters and account for the nonvoluntarism intuition.

---

[171] Often called 'the involuntarism intuition' in the literature. I prefer the 'nonvoluntarism' to the 'involuntarism' label because it does not invite the thought that one believes against his will, as talk of 'involuntarism' does.

Just to provide a first glimpse of the core of habits-endorsement expressivism for the anxious, it will suggest that justification attributions of the form 'S justifiedly believes that p' express that:

(a) S believes that p and

(b) Endorsement of the habits of belief-fixation in virtue of which S believes that p where these habits employ and are constrained by certain epistemic norms –even if the attributor has no clear, nuanced idea of what these epistemic habits and norms are or involve.

This provides a dense introduction of how the dialectic of the argument will unfold and offers a forward glimpse to our positive proposal. In schematic form, this is how the argument of the chapter will go. In section 7.2 I briefly introduce Gibbard's (2003, 2008) plan-reliance expressivism and explain its problems when extended to epistemic justification discourse. In section 7.3, I introduce the positive proposal in some detail and then in section 7.4 explain how the proposed habits-endorsement expressivism can both account for what seems attractive in these two theories and, at the same time, evade the identified problems of plan-reliance expressivism and thus come out as the most explanatorily fruitful. In section 7.5 I add some more touches to our sketch of habits-endorsement expressivism and in 7.6 quickly consider two anticipated objections and gesture towards directions possible answers might go. Finally, in section 7.7 I conclude and preview the argument of the chapter.

Let us now turn to the introduction of plan-reliance expressivism and the identification of its problems.

7.2 Two Problems of Plan-Reliance Expressivism

Roughly, Gibbard's (2003, 2008) plan-reliance expressivism for normative (moral and knowledge) discourses suggests that normative claims like assertions and attributions are 'plan-laden', that is, express 'contingency plans'. Contingency plans are 'determinations' about what to do and what to believe for expected or merely hypothetical scenarios, as Gibbard (2003:53) says. Idealizing somewhat, Gibbard (2003:53-5) also suggests that 'a maximal contingency plan' (or 'hyperplan') is a plan about what to do or what to believe in every conceivable circumstance.

In particular, for moral discourse he suggests that we express plans about the

thing to do. For example, to say that 'Murder is wrong' is to express your plan not to murder people. Or to say that 'Caesar did the right the thing when he crossed the Rubicon' is to express your plan to do the same if the circumstances ever present the same (or analogous) challenges. In later work, he (2008: 16-7) adds some more touches on this account and suggests that moral discourse does not just express plans about what to do but also sentiments. Plans are expressed with some 'emotional flavour', as he puts it. To say, for example, that an act is wrong is to express your plan not to engage in such acts and feel that the act merits resentment by others and guilt on behalf of the perpetrator.

In the case of knowledge discourse, things are slightly different as Gibbard becomes more specific about the kind of plans we express. He suggests that knowledge discourse expresses plans of reliance. Gibbard (2003:227) suggests that knowledge assertions and attributions express noncognitive plans of reliance on the belief in question. For instance the knowledge assertion, 'I know that p' expresses a noncognitive plan to rely on this belief as if I put trust on it. Again, knowledge attributions like 'S knows that p' express a noncognitive plan to rely on S's judgement, namely, the belief that p as if p is something that I put trust on.

Now, initially, one might naturally think that this plan-reliance expressivist theory could extend to cover epistemic justification discourse as well. This would mean that justification assertions and attributions express plans of reliance. For example, to say that 'S justifiably believes that p' is to express your plan to rely on S' judgement that p if circumstances ever present the chance. You plan to rely on p as if you put trust on it and allow it to inform your actions. That is, if the circumstances ever present the chance, you can use p as a means-belief to the satisfaction of your end-desires[172].

Such a transposition of plan-reliance expressivism from knowledge to epistemic justification discourse might prima facie seem promising as, like norm-expressivism we have encountered in Ch.6, plan-reliance discourse has certain attractions as an expressivist theory. On the one hand, as an expressivist story it can avoid the referentialists' problems in a way analogous with our 'toy' norm-expressivist theory in Ch.6. I need not expand on how it can avoid the referentialists' problems and be dull and repetitive, as it is easy to follow how in an analogous way plan-reliance

---

[172] I assume here a mainstream Humean approach to action as it is found in M.Smith (1994) and D.Davidson (2001).

expressivism can accomplish that.

On the other hand, it explains 'the reliance intuition', that is, the strong intuition that normative sentences express attitudes of reliance. Normative sentences like 'p is good or rational' or 'p is justified' or 'I know that p' seem to express, as Austin (1961) has pointed out, trust and guarantee for this claim as if the speaker stakes his reputation for that claim. The speaker expresses his trust to this course of action or certain belief and suggests that he intends to rely on this possible course of action or belief if the circumstances ever demand it.

But despite its attractions, plan-reliance expressivism as an approach to epistemic justifiation discourse seems to have problems we have already foreshadowed in the introductory section 7.1. To come to the point, according to Gibbard (2003) knowledge assertions and attributions express attitudes of plan-reliance, that is, plans to rely on the belief in question and the extension of this account to epistemic justification discourse seems to run into a serious problem from the start.

The problem is this: trivially, given that epistemic justification entails fallibility, justification attribution discourse includes attribution of justified belief even in cases where we think the belief is false. But then, intuitively, if we are rational we won't plan to rely on this belief and this seems a clear *counterexample* to the plan-reliance analysis of justification discourse. If this is the case, then the plan-reliance understanding of justification attributions cannot make sense of this important segment of justification discourse. Let me elaborate in more detail.

While knowledge assertions and attributions are factive and entail truth and non-falsity, justification assertions and attributions are fallible and do not entail truth and non-falsity[173]. You may very well attribute a justified belief to someone and still think that the belief is false while it seems counterintuitive to attribute knowledge and still think the belief is false. For if you believe that the belief in question is false then you won't attribute knowledge in the first place. You might of course attribute knowledge claim like 'S claims to know that p' but this is a different matter as we can attribute knowledge claims and still think that the belief in question is false and S

---

[173] The different constraints on knowledge and justification discourses are nicely spelled out in K.Kappel's (2010; ms) distinction between knowledge-norms (i.e. 'k-norms') and justification-norms (i.e. 'j-norms'). Also, most epistemologists think that knowledge although entails truth is fallible e.g. T.Williamson (2000), J.Greco (2003), D.Pritchard (2010). If this is the case then this could be seen as a problem for fallibilism about knowledge. At any event, this is beyond the point in discussion here.

doesn't really know.

The problem that impedes extension of plan-reliance from knowledge discourse to epistemic justification discourse is straightforward. The problem is that even if we grant the controversial assumption that plan-reliance expressivism is the right expressivist theory for knowledge discourse, clearly, it cannot successfully carry over to justification discourse. For, unlike knowledge attributions, we can easily attribute justified belief but not plan to rely on it because we think is false. At least if we are rational agents, we won't plan to rely on a belief we think is false.

For example, think of the stranded tourist who asks a local for the whereabouts of the museum but the informant intentionally lies to the tourist because he is a sadistic person that maliciously takes pleasure in others people trouble. Now, if we approve of the epistemic practice of asking seemingly trustworthy agents for information but still know that in this case the luckless stranded tourist has been misinformed, we would be inclined to attribute justified belief to the tourist but don't plan to rely on it because we know is false. We would have said something like 'S justifiedly believes that p but still p is false'.

Such cases of justified belief attribution constitute direct counterexamples to any attempt to extend plan-reliance expressivism to epistemic justification discourse. Thus, even if plan-reliance expressivism works for knowledge discourse –which is far from obvious that it does- it can't be transposed to epistemic justification discourse for sure. It can't explain our epistemic practice of attributing justified belief that, nonetheless, we take to be false. If this diagnosis is right, then we need an expressivist theory of epistemic justification discourse that can account for our epistemic practice of attributing justifyied belief but not planning to rely on it because we also think that the belief is false. Plan-reliance expressivism won't do.

As if this problem wasn't enough, plan-reliance expressivism owes us a story about 'the nonvoluntarism intuition'. That is, as any other theory of epistemic justification, it owes us a story of how (justified) belief-fixation is not directly voluntary. The nonvoluntarism intuition is widely held among epistemologists and, yet, norm-expressivism says nothing about how to account for this strong intuition[174]. Let me explain. The nonvoluntarism intuition suggests that we can't voluntarily

---

[174] For the nonvoluntarism intuition see B.Williams (1973), R.Feldman (2001b), R.Audi (2001), D.Owens (2003), W.Alston (2005) and D.Hume (1986).

come to form (justified) beliefs. We can't form justified beliefs or, indeed, any other beliefs at will as belief-fixation seems to be directly nonvoluntary. We *find* ourselves either forming beliefs or with already unconsciously formed beliefs as we don't choose in any direct sense what to believe.

That belief-fixation is directly nonvoluntary can be easily illustrated with a simple, instantaneous psychological 'experiment'. Just try, intentionally, to *choose* to form a novel belief that you have no clue about its truth-value like, for instance, 'There are exactly 100 species of sharks' or a novel belief you think is unjustified or even plain false, say, 'The capital of Nepal is Edinburgh'. Why this novel belief should be one that we have no clue about its truth-value or one we think is unjustified or even plain false will be explained in a moment.

If I may generalize from how this little 'experiment' strikes me, it seems that, sincerely, no matter how hard we try, we can't form these beliefs. We can't voluntarily come to the mental state of believing these propositions. Even if there is a prize or bonus on the task of really believing such propositions and, therefore, we desperately want to form the belief and try to motivate ourselves to do so in order to win the prize, it seems that we can't. If this diagnosis is right, then the philosophical moral of our little psychological experiment is that belief-fixation is not directly voluntary. It just seems that for what human (epistemic) psychology allows belief-fixation is not open to direct choice.

At this point someone might object that this conclusion is a bit too quick as there are at least some cases where we do directly choose what to believe[175]. Possible candidates for directly voluntary belief-fixation might be cases were we choose to form a novel belief that we have just discovered that it is well supported by evidence or plain true. Suppose, for instance, that we just read in an acclaimed encyclopaedia that 'The capital of Nepal is Kathmandu' and then choose to form this novel belief. Accordingly, one might think, in such cases we come to form the novel belief we have chosen to form.

Yet, such cases should not mislead us into thinking that we can sometimes choose what to believe because in such cases we still don't form the novel belief at will, that is, because we tried to choose to form the belief. Rather we form the belief because it

---

[175] See, for example, Carl Ginet (2001).

was considered to bear positive epistemic status, that is, being true, justified or other. Indeed, we would expect agents to form the belief *even before* they find the short time to entertain the thought of choosing to endorse this belief. On the assumption that the agent paid some attention to this piece of information and did not just read it mechanically or absent-mindedly, the belief will simply 'assail' him and be formed nonvoluntarily.

Overall, what such cases really show is not that we can form justified beliefs at will, but only that we seem to nonvoluntarily form beliefs that are found to bear positive epistemic status (justified, rational, evidence-based, true etc.) and shun beliefs that are found to bear negative epistemic status (unjustified, irrational, false etc.). This is why I asked in the little psychological experiment above that we choose to believe a novel proposition that we are totally in the dark about its truth-value or one that we think is unjustified or plain false.

The reason is that our doxastic will is attracted by epistemic standings with positive status and trying to choose to believe a proposition with positive epistemic status might have misled us into thinking that this is possible in some cases. Thus, as we nonvoluntarily form beliefs we find to have positive epistemic status, such cases not only don't come as counterexamples to the nonvoluntarism intuition but also back the intuition that (justified) belief-fixation seems to be not directly voluntary.

Pascal's famous wager illustrates this very point nicely. If you are familiar with Pascal's wager, Pascal urges people who find his pragmatic argument for believing in God convincing to attend masses, say prayers etc. He urges people to immerse themselves in religious practices because he knows very well that belief does not come at will and, therefore, as a means to gradual inducement of religious belief they should immerse themselves in religious practices. In this oblique way, Pascal thinks, agents can come to believe in God's existence due to his pragmatic argument from maximum expected utility for believing in God's existence[176].

Instead, if Pascal's argument had been an epistemic argument, that is, an argument that aims at truth (or any other positive epistemic standing) then Pascal wouldn't have worried because, if cogent, rational agents should nonvoluntarily come to form the belief that God exists. But just because his argument is a pragmatic

---

[176] For an excellent reconstruction and criticism of Pascal's wager see E.Sober and A.Mougin (1994).

one arguing for believing in God's existence from the prudential grounds of maximum expected utility, he knows that agents must somehow manipulate their doxastic will into believing in God's existence.

This much being said about 'the nonvoluntarism intuition', the problem now for plan-reliance is that, in defiance of the nonvoluntarism intuition, as it extends from moral discourse to the epistemic discourse, also implies *uncritical extension* of the voluntarist commitment that we can choose what to believe as we can choose what to do. A mistake that is not entirely uncommon for moral philosophers who are sometimes quick to overlook this important disparity between moral and epistemic discourse[177].

But in the practical domain this voluntarist commitment is relatively unproblematic as we literally have the power to make practical choices and act at will; at least if we are rational enough to be able to stick to our best choice. Just try out a practical version of our little psychological experiment and choose, say, to raise your hand or scratch your head. This ability for practical choice seems to simply follow from our profound ability to somehow be free agents, that is, agents with the marvellous ability to freely choose among alternative courses of action. This is, after all, why we often say things like 'p is the right course of action and this is what I choose to do'.

But as we have just seen, in the epistemic domain this voluntarist commitment seems puzzling, as it appears that we can't choose what to believe. We can't choose at will what to believe, as doxastic volition is not under our direct voluntary control. Even if we want to believe at will and form the intention to believe at will, this appears psychologically impossible; at least, for what human psychology allows. This is why saying things like 'p is the right thing to believe and this is what I choose to believe' strike us as unnatural in terms of ordinary discourse. Rather, it is more natural to say that 'p is the rational thing to believe and this is what I believe' because we tend to nonvoluntarily form beliefs we take to have positive epistemic status (evidence-based, rational, justified etc.) and shun beliefs we take to have negative epistemic status (not evidence-based, irrational, unjustified etc.).

Unfortunately, though, plan-reliance expressivism defies the nonvoluntarism

---

[177] See for example C.Korsgaard (1996: 92-3)

intuition and suggests that we can form epistemic plans about what to believe and rely on at will. It uncritically extends to cover both moral and knowledge discourse and this presupposes that we can form beliefs at will, as if beliefs were a matter of our direct voluntary control. This seems to be especially the case because the notion of 'planning' has strong voluntary practical connotations. We can, intentionally, choose among various practical plans about what to do in certain circumstances. As I graduate I can choose, say, to plan to pursue a Ph.D rather than plan to pursue my luck in the job market.

Yet, although planning may seem relatively unproblematic in the practical domain where we do seem to be able to choose to form plans about what to do among alternatives, in the epistemic domain planning seems clearly problematic. We don't seem psychologically able to form at will epistemic plans about what to believe and rely on. Even if I form the intention to form the epistemic plan to only believe and rely on what is, let us say, coherent with my overall web of beliefs this does not show that I will eventually be able to form such a coherentist epistemic plan and come to believe and rely on what is coherent with my web of beliefs.

A fortiori, even if I somehow manage to follow my intention and come to form the coherentist epistemic plan, it is far from obvious that I will be able to form beliefs according to the dictates of the epistemic plan. Thus, intuitively, we can't exert any direct voluntary control on our doxastic volition and thereby to belief-fixation.

The problem for plan-reliance expressivism is that it seems to assume that as in the practical sphere of our agential life we can choose what practical plans to adopt for the regulation of our practical conduct, equally in the epistemic sphere of our agential life we can choose what epistemic plans to adopt for the regulation of our epistemic conduct. But as the simple psychological experiment above has illustrated, this uncritical extension of the voluntarist commitment from the practical/moral domain to the epistemic it is far from being unproblematic. For, belief-fixation seems to be directly nonvoluntary and plan-reliance seems to assume again that we can choose what to believe and rely on through choosing epistemic plans.

This concludes a concise introduction of Gibbard's plan-reliance expressivism and the identified problems it runs into if extended to cover epistemic justification discourse. The scene is now ready for the advent of habits-endorsement

expressivism, which is meant to constructively synthesize insights in an original manner that will allow the theory to assert itself as the explanatorily more fruitful. Let us now turn to the exploration of this conceptual synthesis.

7.3 The Possibility of Habits-Endorsement Expressivism

So far, we have argued that plan-reliance expressivism defies the nonvoluntarism intuition about epistemic justification discourse and, in addition, runs into counterexamples because it fails to explain how we can attribute justified belief and still think that the belief is false and, thereby, as rational agents don't plan to rely on it.

In this section we constructively synthesize insights in order to build an original norm-expressivist theory for epistemic justification discourse that keeps what is attractive in both norm-expressivism and plan-reliance expressivism theories while avoids the identified problems of plan-reliance expressivism. This original expressivist theory will be constructed on the basis of the notion of epistemic habits. The rich and versatile notion of habits will be placed at the foundations of the proposal and will allow us to connect and cement insights in an original way that will, hopefully, prove elegant and explanatorily fruitful. Let us now turn to theory building.

Needless to say, in order to analyse what we mean when we make justification assertions and attributions we need to reflect on what we implicate when we attribute justified belief. What we seem to express in such discursive contexts. This is what we do then. But in order to simplify things, in what follows I will focus exclusively on what we mean when we make justified belief attributions of the form 'S justifiedly believes that p', though, the expressivist semantic picture I will paint should be broadly analogous for justification assertions as well.

The first obvious thing to notice is that in discursive contexts where we say 'S justifiedly believes that p' we express that 'S believes that p'. This seems quite easy to follow as expression of attribution of justified belief (i.e S justifiedly believes that p) obviously implicates attribution of belief (i.e. S believes that p). The second obvious thing that quickly comes to the eye is that we express some sort of endorsement or approval for S's believing that p. That is, we seem to see 'S's believing that p' in some favourable light.

For as we have already noted in section 6.2, expressivist approaches reverse the order of semantic explanation and embark from the element of endorsement that the referentialists fail to accommodate due to 'the Moorean/Humean lesson'. And, of course, as expressivist approaches embark from the element of endorsement, habits-endorsement expressivism could not but also embark from the element of endorsement. Accordingly, when we say 'S justifiedly believes that p', we seem to implicate a general attitude of endorsement or approval of 'S's believing that p'.

This much of analysis is pretty standard for an expressivist account of justification. The apparent question now is what sort of endorsement this general attitude of endorsement is. It is here that the notion of epistemic habits enters the theoretical scene in order to play the needed explanatory role. As we all acknowledge from our everyday lives, habits constitute 'the great guide of human life', as Hume (2005:44) has put it[178]. We have sleeping, eating, cooking, drinking, studying, writing, talking, walking, shopping, playing, thinking, driving, cycling, diving, cleaning habits and much more. If therefore Hume is right, as he seems to be, and we are at bottom such 'creatures of habit' then the epistemic segment of our lives couldn't be an exception. Our epistemic lives, and hence, justified belief-fixation should accordingly be guided by epistemic habits, namely, habits of belief-fixation.

Indeed, intuitively, (justified) belief-fixation seems importantly to be habitual, that is, based on habits of belief-fixation. For, as we have seen in section 7.2, belief-fixation seems to be directly nonnvoluntary and for the most time unreflective and unconscious and these are emblematic properties of habits. That is, the very notion of habits as such implies both at least some grip on our psychology and quite often unreflectiveness and unconsciousness. This is why sometimes we say things like 'He is at the mercy of the force of his evil habits' or 'He always buys the same toothpaste out of habit and doesn't even realize!'. Or think of the more concrete example of the habit of going by the stairs rather than the elevator. This habit is a habit insofar it has a grip on our psychology, namely, we are disposed to act accordingly and we do so for at least the most part unreflectively and unconsciously.

As many epistemologists take the lead from the pervasiveness of habits in everyday life and the properties belief-fixation and habits share and from time to

---

[178] To be exact, 'Custom, then, is the great guide of human life'(2005:44).

time talk about 'epistemic habits', 'habits of belief-formation', 'cognitive schemes, plans, stances, policies or heuristics' etc., the idea that epistemic habits could lie at the heart of our epistemic practices should not come as a surprise[179].

Thus, if we would like to build an expressivist story that would capture the habitual character of justified belief-fixation the idea would then be that when we express endorsement of 'S's believing that p', our endorsement is an endorsement of S's epistemic habits, that is, S's habits of belief-fixation in virtue of which S believes that p. In this picture, epistemic habits are bestowed with the cognitive functional role of (justified) belief-fixation, that is, their function in our mental economy is to guide (justified) belief-fixation.

Here is an example of how this is supposed to work. Suppose I say 'Mary justifiedly believes that her next pottery class will take place this weekend'. According to the expressivist semantic picture we have painted so far what I, the attributor, mean with this sentence is that Mary believes that her next pottery class will take place this weekend and endorse Mary's habits of belief-fixation in virtue of which she has formed this belief. Maybe the habit of belief-fixation responsible for the formation of Mary's belief is her habit to trust what reliable sources of information (like the pottery classes website) say.

Of important note is that the attributor may endorse S's habits of belief-fixation in virtue of which S believes that p, even if he has no clear idea of what these epistemic habits are or involve. This seems to occur quite often, if not most of the time, in our epistemic practice of justified belief attribution as attributors are not aware of other agents' epistemic habits but still readily attribute justified belief.

If we take the example above, as an attributor I could still feel inclined to attribute justified belief to Mary, even if I didn't have much of an idea of the habit of belief-fixation in virtue of which she has formed the belief. I could attribute justified belief to Mary just on the basis of, say, that I trust her as an epistemic agent. That this is the case should not perplex or puzzle us, for thinking that an attributor should be aware of S's habits in order to attribute justified belief would seem cognitively too demanding and, as a result, it would have over-intellectualized and thereby distort

---

[179] See D.Hume (1986, 2005), C.S.Peirce (1991), J.McDowell (1995) , L.Zagzebski (1996), R.Wedgwood (2009), G.Harman (1996), A.Goldman (1978), P.Helm (1994). The importance of habits in our lives is also often acknowledged in literature. Compare F.Dostoyevsky (2008): '…to what monstrous extent we are creatures of habit'.

the everyday practice of attributions of justified belief.

Such attributions of justified belief without awareness of the epistemic habits in play is especially widespread in our modern societies of overspecialization and rigorous division of epistemic labour where the epistemic practice of attribution of justified belief to experts is an everyday practice. We say, for instance, 'The physicists' community justifiedly believes that the universe is constituted by minute particles called atoms' and, according to the proposal, express endorsement for the habits of belief-fixation in virtue of which the academic community of physicists has reached this conclusion, even though the vast majority of us have no clear idea what these habits are.

Nowadays, perhaps, most laymen have some inkling that science goes hand with hand with experimenting, as it is now common lore that science progresses by means of experimental trial and error. But for most people this general inkling is all they are aware of scientists' habits of belief-fixation, as they are far from being experts in the history and philosophy of science.

Clarifying, though, what sort of endorsement the attributor's endorsement of 'S's believing that p' is seems to be just the tip of the (semantic) iceberg as there is yet much more going on when we express such endorsement. The third thing to notice is that in such discursive contexts the habits of belief-fixation we endorse seem to implicate *constraint* by certain epistemic norms. Reasonably, when we attribute justified belief and express endorsement for S's habits of belief-fixation we also imply that these habits employ and are constrained by certain epistemic norms.

For, surely, by S's epistemic habits (no matter what these are) own lights not everything goes as some beliefs count as justified and others as unjustified. Habits of belief-fixation, thus, commit to epistemic norms that license beliefs as justified or rule out beliefs as unjustified. These norms might be coherentist, foundationalist, deontologist, reliabilist, virtue-theoretic, relaxed or austere, conservative or revisionist etc.

If we take the Mary example above, where 'Mary justifiedly believes that her next pottery class will take place this weekend' because she is in the habit to trust what reliable sources of information (like the pottery classes website) say, we may say that her habit employs and is constrained by certain epistemic norms about testimony.

That is, norms dictating what sort of testimonial sources of information are reliable and what sources are not. In this case her norms seem to be 'credulist' as they are relaxed and permissive enough, in the absence of any countervailing reasons, to take the pottery classes' website to be a reliable source of information.

Again, as with epistemic habits of belief-fixation, the attributor might have no clear idea of what these epistemic norms employed by the habits are or involve. Attributors can readily ascribe justified belief to Mary but neither have a clear idea of Mary's habits of belief-fixation nor of the employed norms. All that is needed for attribution of justified belief is endorsement or approval of the habits of belief-fixation in virtue of which Mary formed the belief. That said, our rather simple 'Mary example' should not give the wrong signals, namely, that the relation between an agent's habits and employed norms is all that simple and monolithic. On the contrary, the relation between habits and norms is very complex and I will return to this issue in section 7.6. But for the time being our priority is to outline the basics of the theory and we have to stay focussed on that.

Let us now take stock of where we have arrived at so far. According to habits-endorsement expressivism, attributions of justified belief like 'S justifiably believes that p' express that:

(a) S believes that p and

(b) Endorsement of the habits of belief-fixation in virtue of which S believes that p where these habits of belief-fixation employ and are constrained by certain epistemic norms. Awareness of what these epistemic habits and norms are or involve is not necessary for the attributor.

But it seems that there is still plenty that is implicated in such contexts, as endorsement of habits of belief-fixation seems to implicate and be *constrained* by more. This leads to two more points: a fourth and a fifth one. Fourth, as a normative discourse, in justification discourse we tend to have *attitudes of appraisal* towards the agents[180]. We implicate praise for people that have justified beliefs because this is a cognitive achievement and, accordingly, implicate blame for people that have unjustified beliefs because this is a cognitive failure. We say things like 'Copernicus justifiedly believed in heliocentrism well ahead of his contemporaries' and imply that

---

[180] See J.Greco (2003) for appraisal attitudes expressed in knowledge discourse.

Copernicus is to take credit for having this justified belief or say things like 'Many people unjustifiably believe in astrological predictions' and imply that these people are blameworthy for having this unjustified belief.

Yet, we don't *always* imply praise when we attribute justified belief as there are cases where we think that, although 'S justifiedly believes that p', there isn't much to praise because the belief was all too easy to acquire. There are cases where an agent has a justified belief due to, let us say, perceptual evidence but we don't praise or give credit to the agent because acquiring this justified belief was all too easy and can't be seen as a cognitive achievement. Think, for example, of the agent who thinks that 'ripe bananas are yellow' because he has seen a lot of ripe bananas and all were yellow. Inductive scepticism aside, this is a case where we are inclined to attribute justified belief but no praise, though, we also don't want to blame people for having a justified belief even if this was acquired 'on the cheap' and does not count as a cognitive achievement.

In the light of such cases of justified belief attribution where we don't imply praise, we should avoid commitment to the thesis that justified belief attribution always conveys praise. What we should say instead is that justified belief attribution *allows* praising and forbids blaming people for having justified belief. For although justified belief attribution seems to allow praise it doesn't necessitate praise and, of course, always forbids blame. It always forbids blame because even on cases of justified belief attribution where the belief was acquired on the cheap we wouldn't by any means be inclined to attribute blame.

Fifth, as J.Austin (1961) and Gibbard (2003) have pointed out, epistemic discourse also implicates attitudes of trust and reliance. We *trust* the habits of belief-fixation we endorse and therefore *rely* on their belief output, even when we have no clear idea of what these habits of belief-fixation involve. For example, think again of the epistemic practice of attribution of justified belief to experts. We say, for instance, 'Mathematicians justifiably believe that π is approximately 22/7 and an irrational number' and express endorsement for the habits of belief-fixation in virtue of which the academic community of mathematicians has reached this conclusion. We trust that these habits of belief-fixation are truth-conducive and thereby rely on their belief output, even if we have no very clear idea of what these habits involve.

Now, if endorsement of habits of belief-fixation implicates and is constrained by both appraisal attitudes and attitudes of trust and reliance, then our analysis should somehow account for that. Otherwise, our analysis would appear elliptical in these central aspects of justification discourse. And to account for these central aspects our analysis needs to incorporate in the initial scheme the attitudes of appraisal and trust/reliance. The obvious way to do that is to suggest that condition (b), namely, 'endorsement of habits of belief-fixation in virtue of which S believes that p' rationally commits to a regulative norm which:

(i)   Permits praise and forbids blame for agents whose beliefs are formed in virtue of such habits of belief-fixation.

(ii)   Requires trust on these habits of belief-fixation and reliance on their belief output.

I say here 'rationally commits' to a regulative norm because we expect all rational agents who attribute justified belief, if made explicit to them, to acknowledge commitment to attitudes of appraisal and trust/reliance. If an agent who attributes justified belief and by the lights of our analysis endorses S's habits of belief-fixation, refuses to acknowledge commitment to a regulative norm of attitudes of appraisal and trust/reliance we would be inclined to think that this agent is missing something. Perhaps he is conceptually confused and needs some more explaining of what this regulative norm is about and why as an attributor of justified belief he is committed to that. But if he shows signs of clear understanding of what is the point at issue and yet refuses to acknowledge commitment to this regulative norm, then we would be inclined to coin him irrational and abandon the futile argument with him.

So far, so good, but still condition b (ii) seems too crude as it stands. For, although we trust the habits of belief-fixation we approve of, there are cases of justified belief attribution where as attributors we do not rely on their belief output in spite of all our trust and approval. There are cases where although we approve of S's habits of belief fixation we have reasons to think that the produced justified beliefs are likely to be false or even be plain false. These are the cases where we say things like: 'S justifiedly believes that p but I still think that p is false'. If this is right, as it seems to be, then reliance on the habits' belief output should only be *defeasible* as there are cases where agents may have good reasons to think that despite attribution

of justified belief and, thereby, trust and endorsement of S's habits of belief-fixation they should not rely on their belief output. That is, the beliefs these habits of belief-fixation produce.

Think, for instance, of the police detective that arrives at a certain belief about who is the murderer. I, as an attributor, I am inclined to attribute justified belief to him and say that 'S justifiedly believes that Mary is the murderer' but still think that the belief is false indeed and therefore don't rely on this belief. This might happen because, say, I have reliable testimony that Mary has a strong alibi and therefore all evidence against her is merely circumstantial. Yet, I attribute justified belief to the detective because I endorse and trust his habits of belief-fixation in virtue of which he has reached this belief. I may trust and endorse his habits of belief-fixation for a variety of reasons, perhaps, because he has the reputation of being an excellent detective with difficult crimes resolved in his record etc. We can thus talk of *default reliance* on the belief-output of habits of belief-fixation where reliance is *defeasible* and not mere *reliance simpliciter*.

Interestingly, someone might naturally suspect that the converse could also obtain, namely, that as we trust habits of belief-fixation and only defeasibly rely on their belief output we could, conversely, rely (simpliciter or defeasibly) on the belief output of certain habits of belief fixation that we only defeasibly trust and rely on them.

This could happen only if there are cases of justified belief attribution where as attributors we invest only 'default trust' on the habits of belief-fixation we endorse because we have reasons to think that the subject's habits are not ideal from an epistemic point of view. We have reasons to think that the habits of belief-fixation the subject is relying on are not the ones he *should* be relying on, even if a certain belief produced by these habits of belief-fixation is considered to be justified. And this, of course, could happen only in cases where attributors have at least some general idea of what the subject's epistemic habits are or involve.

Yet, this suspicion is too quick because in such cases the attributor would not say 'S justifiedly believes that p'. Instead, he would say things like: 'S's belief that p is justified, though, the habits of belief-fixation responsible for this belief are not ideal'. The difference between the two attributions is subtle but important because what we

attribute to the subject in such a case of 'default trust' on the subject's habits of belief-fixation is *having a justified belief* and not *justified believing* and this, accordingly, should not be included in an analysis of 'S justifiedly believes that p'.

The attribution is of having a justified belief not of justified believing because the subject's habits of belief-fixation responsible for the fixation of this belief are not ideal from an epistemic point of view. The subject has a justified belief but this belief was not acquired by means of epistemic habits we approve of. This is why the attributor says ''S's belief that p is justified, though, the habits of belief-fixation responsible for this belief are not ideal' instead of 'S justifiedly believes that p'.

Here is an example. Think of a German subject during WWII who believes that 'French resistance has collapsed' because he has the habit of belief-fixation to uncritically accept whatever people in authority say and by means of this habit has formed the aforementioned belief because it was officially proclaimed by Goebbels, the notorious Nazi minister of propaganda. In this case, as an attributor if I somehow diagnose that S has this habit of belief-fixation and that this habit is responsible for the fixation of the -indeed justified- belief 'that the French resistance has collapsed', then I would be inclined to say something like 'S's belief that p is justified, though, the habit responsible for the fixation of this belief is not ideal'. But I wouldn't be inclined to say 'S justifiedly believes that p' because that would imply that he has acquired the belief by means of a truth-conducive habit and this was not the case.

Such cases of attribution of having a justified belief would have also created problem for condition b (i), namely, that endorsement of habits of belief-fixation 'permits praise and forbids blame'. It would have created problem for b (i) because in such cases of attribution of having a justified belief the attributor also conveys blame because of cognitive failure. The subject acquired the justified belief on the basis of a habit that is not truth-conducive and could have easily produced an unjustified belief instead. But conveying blame contradicts b (i) that clearly forbids blaming and this would have constituted a counterexample to b (i). Fortunately, as we said, we need not worry about this because what we are analysing here is the practice of justified belief attribution and not the practice of attribution of having a justified belief.

This point set to the one side, the last semantic observation is that even this trust

on habits of belief-fixation and 'default reliance' on their belief output seems to rationally commit to other epistemic habits, this time *habits of reliance*. That is, a habit to rely on certain habits of belief-fixation and a habit to rely on their produced beliefs. For, as habits are acquired through constant repetition and, indeed, habituation into a practice, reasonably, it is to be expected that trusting certain habits of belief-fixation and relying on their belief output over a substantial period of time will create habits of reliance on these habits of belief-fixation and their belief output.

In this sense, trust on habits of belief fixation and 'default reliance' on their belief output commit themselves to habits of reliance, that is, a habit to trust and rely on certain habits of belief-fixation and a habit to rely on their belief output. But the habits of reliance on belief output should also be defeasible because, as have just seen above, there are cases where we do not rely on the belief output of the habits we approve of. Thus, habits enter the picture of habits-endorsement expressivism twice: first, as habits of belief-fixation and, second, as habits of reliance on these habits of belief-fixation and their belief output.

Let us now pull the strings of our enriched analysis of condition (b), namely, of 'endorsement of habits of belief-fixation in virtue of which S believes that p'. In the footsteps of our diagnosis above of what justified belief attribution implies, we analyse condition (b) in terms of rational commitment to a regulative norm which:

(i)   Permits praise and forbids blame for people whose beliefs are formed in virtue of such habits of belief-fixation.

(ii)   Requires trust on these habits of belief-fixation and 'default reliance' on their belief output. Trust and 'default reliance' also implicate habits of reliance on these habits of belief-fixation and their belief output where the habit of reliance on the belief-output of the trusted habits is only defeasible.

We are now poised to bring all the pieces of our analysis together and see what the big picture is like. According to habits-endorsement expressivism, attributions of justified belief like 'S justifiably believes that p' express that:

(a) S believes that p and

(b) Endorsement of the habits of belief-fixation in virtue of which S believes that p where these habits of belief-fixation employ and are constrained by certain epistemic norms. Awareness of what these epistemic habits and norms are or involve

is not necessary for the attributor.

Condition (b) is then further analysed in terms of rational commitment to a regulative norm which:

(i)   Permits praise and forbids blame for people whose beliefs are formed in virtue of such habits of belief-fixation.

(ii)   Requires trust on these habits of belief-fixation and 'default reliance' on their belief output. Trust and 'default reliance' also implicate habits of reliance on these habits of belief-fixation and their belief output where the habit of reliance on the belief-output of the trusted habits is only defeasible.

This concludes our sketchy introduction of habits-endorsement expressivism. With habits-endorsement expressivism now introduced let us turn to the next section and see how it fares against the identified problems of plan-reliance expressivism.

7.4 The Explanatory Power of Habits-Endorsement Expressivism

Let us now examine how this proposal, first, captures what is attractive in norm-expressivism and plan-reliance expressivism theories and, second, tackles the identified problems of plan-reliance expressivism and therefore asserts itself as the explanatorily more fruitful.

First of all, as an expressivist story habits-endorsement expressivism dodges the referentialists' problems in a way analogous to our 'toy' version of norm-expressivism in Ch.6. As I have already said before, there is no reason to be dull and repetitive, as it is quite easy to follow how it does that. Second, habits-endorsement expressivism easily captures the attractive intuitions that, respectively, norm-expressivism and plan-reliance expressivism capture because these are virtually build into the account. Plus, it explains what I call 'the habits-constraining intuition'. Let me elaborate.

It can easily explain 'the endorsement intuition' because, like other expressivist stories, it starts from what agents endorse when they express normative sentences. For habits-endorsement expressivism, it expresses endorsement for the habits of belief-fixation in virtue of which the belief was formed. Thus, endorsement is pretty much built into the account. It also accounts for 'the norm-constraining intuition' because for the proposal epistemic justification sentences express endorsement of habits of belief-fixation where these habits of belief-fixation employ and are

constrained by certain epistemic norms. Again, constraint by epistemic norms is pretty much built into the account.

Finally, it explains 'the reliance intuition' because for habits-endorsement expressivism, expression of endorsement of habits of belief-fixation rationally commits to a regulative norm that requires trust on these habits of belief-fixation and 'default reliance' on their belief output. Indeed, we come to have habits of reliance on our habits of belief-fixation and their belief output. Once again, reliance is build into the account and therefore habits-endorsement expressivism easily captures 'the reliance intuition' too.

Third, habits-endorsement expressivism explains 'the habits-constraining intuition', namely, the intuition that belief-fixation is constrained by the operation of habits of belief-fixation – an intuition that as we have mentioned above has been attested by many epistemologists. It explains 'the habits-constraining intuition' because for the proposal epistemic justification assertion and attributions express endorsement of certain habits of belief-fixation.

Fourth, habits-endorsement expressivism outflanks the problems we identified for plan-reliance expressivism. It evades the counterexamples plan-reliance expressivism runs into and accounts for the nonvoluntarism intuition. Let us start expounding how it can do this much with how it easily evades the counterexamples plan-reliance expressivism runs into.

Habits-endorsement expressivism evades the counterexamples plan-reliance expressivism runs into because it easily accounts for attribution of justified belief in cases where we think that nonetheless the belief is false and therefore don't plan to rely on it. We can do that because we can endorse S's habits of belief-fixation even though we might think that his habits of belief-fixation in this particular case have led him astray to a false belief and, therefore, do not intend or plan to rely on this belief if circumstances may demand it.

For example, think again of 'the stranded tourist example' where the tourist asks a local for the whereabouts of the museum but the informant intentionally lies to the tourist because he is a sadistic person that maliciously takes pleasure in others people trouble. In such a case, although we know that the tourist has been misdirected, we could still attribute justified belief to the tourist because we approve of the habit of

belief-fixation of asking seemingly trustworthy locals for information, though, we wouldn't intend or plan to rely on the belief. We would say something like 'S jusifiedly believes that p, though, p is false'. This is exactly why we have been careful enough to say that reliance on the belief-output of our endorsed habits of belief-fixation is only defeasible and, hence, 'default'. For attribution of justified beliefs should be able to handle such cases, something that plan-reliance as applied to epistemic justification discourse does not.

Let us now turn to how the proposal accounts for the nonvoluntarism intuition. Habits-endorsement expressivism can account for the nonvoluntarism intuition because it suggests that we nonvoluntarily form (justified) beliefs due to habits of belief-fixation. The gambit is made by the introduction of the theoretically versatile notion of habits. For as we briefly mentioned in section 7.4, the very notion of a habit implies a certain grip on our psychology and for the most part unreflectiveness and unconsciousness, and these properties of habits entitle us to account for the directly nonvoluntary character of belief-fixation.

These properties of habits entitle us to do so because by appeal to these properties, namely, a grip on our psychology, unreflectiveness and uncosciousness we can suggest that (justified) belief fixation is not directly voluntary due to the operation of epistemic habits. Belief-fixation is not directly voluntary because is guided by habits of belief-fixation and 'a grip on our psychology' is one of the very properties habits as such exhibit. This is why we often say 'Oh, I struggle with this bad habit but to no avail. It seems to have taken the best of me'. Thus, the nonvoluntary grip habits have on our psychology explains why our justified belief-fixation is also directly nonvoluntary. Justified belief-fixation is directly nonvoluntary because it is guided by habits of belief-fixation and these by nature have a grip on our psychology.

Habits-endorsement expressivism also accounts for epistemic improvement, something we briefly touched in Ch.6 with our distinction between 'the realist existential assertion'/'the minimalist existential assertion' that entitled us to talk of 'minimally justified beliefs'. For it allows that through an 'ideal epistemic observer standpoint' we can reflect on and criticize our habits of belief-fixation and come to have different sort of habits by initiating cultivation of the habits of belief fixation

we think ideal[181]. Such ideal habits of belief-fixation are, presumably, to be habits that are conducive to the social good of truth and are therefore approved and encouraged from the social point of view.

Thus, through the ideal epistemic point of view the agent can criticize and change his habits of belief-fixation, even though this can often prove a hard and psychologically demanding process if the habits are deeply-rooted in our psychological profile. The agent can change his habits of belief-fixation with more truth-conducive habits and therefore his justified beliefs come out of cognitive ability. They come out of his ability to observe his epistemic habits, reflect on them and if not satisfied with them criticize and revise them with epistemic habits he considers better from an epistemic point of view.

This process of self-observation, criticism and revision of epistemic habits occurs because there are cases where we have reasons to think that our habits are not ideal from an epistemic point of view. We have reasons to think that the habits of belief-fixation we are relying on are not the ones we *should* be relying on. This could happen in cases where agents have at least some general idea of what their epistemic habits are or involve, as they might grow dissatisfied with these habits and decide to revise them with others they consider better from an epistemic point of view.

Here is an example of how such habits-revision is to work. John is a great admirer of crime fiction and especially of the all-time-classics of A.Conan Doyle's Sherlock Holmes and A.Christie's Heracles Poirot. He thinks that these fictional characters embody the ideal epistemic agents and aspires to emulate their habits of justified belief-fixation. So, after some meticulous self-observation of his habits of belief-fixation he becomes convinced that his habits are far from being even similar to his heroes and decides to revise them. He finds, say, that he is in the habit of being impatient and jumping to conclusions, lacks persistence, he is often inattentive and overlooks facts, prone to probabilistic fallacies, easily clouded by selfish desires, emotions etc. while his epistemic heroes are in the habit of being patient, thoughtful, persistent, attentive and diligent, resist probabilistic fallacies and being clouded by their selfish desires, emotions etc.

Our friend John, thus, decides to take measures and start cultivating the epistemic

---

[181] 'Ideal observer' ideas can be traced back to A.Smith (2007) and I.Kant (2006). Such ideas are often explored by contemporary philosophers, for example, M.Smith (1994) and R.Firth (2007).

habits of his ideal epistemic heroes Sherlock and Heracles. Every time he thinks what sort of epistemic habits should have he asks: 'What epistemic habits Sherlock and Heracles have?' and then tries to emulate these habits and cultivate them in order to gradually through repetition come to get a grip on his psychology. He learns, for instance, to resist his probabilistic fallacies that, in general, as epistemic agents we are so psychologically prone to fall prey to (as abundant evidence from cognitive psychology suggests)[182]. That is, as evidence from cognitive psychology suggests, our intuitions about probabilities often lead us astray as it happens, for example, in 'the conjunction fallacy', 'the gambler's fallacy' etc.

Take for example the conjunction fallacy. The conjunction fallacy is a probabilistic fallacy often committed by many epistemic agents prone to think that a conjunction of probabilities is more probable than one of its conjuncts. This is, of course, mistaken as any individual conjunct is necessarily more probable than the whole conjunction itself. Any P(a) is more probable than P(a) *and* P(b); given of course that the probability of P(b) is not one or zero. Reasonably, it is more probable that it will rain today than both it will rain today and pass my viva. Yet, like John, many agents quite often seem to take this wrongly, presumably, because they mistakenly associate conjunction with the notion of addition and therefore increase of probability. But, unlike addition in arithmetic calculus, addition in probability calculus requires multiplication and, therefore, decreases rather increases the probability rate as it makes the events less likely to all conjunctively occur.

However, an agent like John who aspires to improve his epistemic performance, can either at some point understand how this fallacy works on his own or be instructed from reliable epistemic peers who already understand how the fallacy works. And once he realizes how the fallacy works (and the psychological mechanism of the habit of belief-fixation operating behind it), then he can take steps to get rid of the habit of belief-fixation in virtue of which the false beliefs are formed and learn to resist committing the fallacy. A fortiori, he can through repetition and habituation come to have the habit of belief-fixation to believe that the probability of any conjunct is greater than the probability of the conjunction.

He can start cultivating the epistemic habit to consider a conjunct more probable

---

[182] See A.Tversky and D.Kahneman(1985)

than the probability of the whole conjunction. The way to do that is to pay extra attention and diligence when he encounters conjunctive probabilities, keep an eye on the fallacy and resist the disposition to believe that a conjunction is more probable than an individual conjunct. In time, if he is diligent, persistent and patient enough, he will get rid of the fallacious epistemic habit and come to have the epistemic habit to consider a conjunction of probabilities less probable that any of its conjuncts.

As one might have noticed in 'the John example' just canvassed, habits-revision relies substantially on the character of the agent in question. For as habits by nature have a grip on our psychology, habits-revision might be often a psychologically demanding process and therefore an agent should exhibit some virtues (or excellences) of character if he is to manage to effect habits-revision. That is, in our example learn to resist the habit of belief-fixation responsible for the conjunction fallacy and moreover acquire the habit of belief-fixation that sets things right and take individual conjuncts to be more probable then the conjunction.

Such virtues of character are diligence, patience, persistence, cautiousness, courage, integrity, a love of truth, self-respect etc. This point, moreover, raises interesting virtue-theoretic issues about the role character plays in matters epistemic like the relation between normative habits-formation and character, character formation itself etc. that, unfortunately, I can't pursue and do full justice with them here.

This discussion concludes our sketch of how habits-endorsement expressivism captures the attractions of norm-expressivism and plan-reliance expressivism while at the same time evades the problems we identified for plan-reliance expressivism. Perhaps, it is not as theoretically simple as these theories but this does not disrespect Occam's razor because what habits-endorsement expressivism loses in simplicity it gains in explanatory power and this is the primary desideratum for any theoretical choice.

If we are right then, habits-endorsement expressivism has a good prima facie claim for being explanatorily more fruitful of plan-reliance expressivism. But this presumptive case for habits-endorsement expressivism does not lull us into complacency, as the theory is still crude and sketchy in many key respects and may even fail in some crucial respect yet to be noticed. What is more, there might be

another theoretical story that can reap the explanatory fruit habits-endorsement expressivism does (or even more) in a simpler and more economical way. If that is the case, then that theory will have a better claim for being the best of the lot, but as far as I can see in the theoretical horizon there is no such theory. At any rate, what I am convinced of is that habits-endorsement expressivism is an interesting approach that needs to be further explored in good hope that it can make full sense of the epistemic justification discourse.

With the sketch of habits-endorsement expressivism at hand and with its explanatory sharpness just tested, in the next section I add some more touches to habits-endorsement expressivism. These touches purport to show that habits-endorsement expressivism is not all that simple and monolithic, as our crude sketch in section 7.3 might have suggested.

7.5 Some More on Habits-Endorsement Expressivism

In the section 7.3 I have argued that justified belief attributions (and justification discourse more generally) implicate endorsement of habits of belief-fixation where these habits employ and are constrained by epistemic norms. When I suggested this I also hinted that the relation between epistemic habits and norms should not be assumed that is all that simple and monolithic and promised to return to this point in this next section 7.6. This is what I first take up in this section where I add some more touches to the proposal.

As I suggested that justified belief attributions express endorsement of habits of belief-fixation constrained by certain epistemic norms, I gave 'the Mary example' where 'Mary justifiedly believes that her next pottery class will take place this weekend'. Mary, recall, forms this belief because she is in the habit to trust what reliable sources of information (like the pottery classes website) say and her habit employs and is constrained by certain epistemic norms about testimony. That is, norms dictating what sorts of testimonial sources of information are reliable and what sources are not. In this case, her norms seem to be 'credulist' as they are relaxed and permissive enough, in the absence of countervailing reasons, to take the pottery classes' website to be a reliable source of information.

But of course, the rather simple 'Mary example' should not convey the false impression that agents employ always the same epistemic habits (and epistemic

norms) across all contexts. This might happen in certain cases but surely it doesn't happen always and, perhaps, it doesn't happen often too. For, intuitively, there are plenty of cases where the same agent in different contexts employs different epistemic habits. This is sufficiently obvious from the context-sensitive way justification assertions and attributions (as well as other normative notions like 'good') are being used in ordinary discourse[183].

In ordinary discourse justification assertions and attributions are being used in different contexts with a varying degree of demandingness. What counts as justified belief in one context might not count as justified belief in another context, as the context swiftly changes and the epistemic norms' demandingness varies accordingly. Intuitively, this fluctuation of the demandingness of the epistemic norms that license justified belief happens because our epistemic lives are very complicated and, as a result, different contexts and circumstances may be found to pose different epistemic demands.

Here are two examples where the demandingness of epistemic norms in play varies according to the demands of the context. First, a philosophy student might switch epistemic habits when in the epistemology seminar. So, while in ordinary life and outside the epistemology seminar she has the epistemic habit to believe what she sees and, hence, accepts that she justifiably believes that has hands, when in the seminar the epistemic habit changes, epistemic norms are accordingly conceived in much more demanding terms and she now resists that this belief is really justified. She thinks that mere seeing her hands is not sufficient to really justify the belief that she has hands, as a Cartesian evil demon might be deceiving her.

Second, a doctor might switch epistemic habits from case to case and from patient to patient relative to what practical stakes his justification assertions might imply. He might, for instance, think that he justifiedly believes that in the case of a patient with a harmless flu he should prescribe some antibiotics. Yet in the case of life-threatening situations he might be reluctant to consider himself as a justified believer of what medication should be prescribed. He might be reluctant because any such

---

[183] As I have already canvassed in section 1.4, it might be noticed that epistemic justification is not much used in ordinary discourse as it is more of an academic term. This is I think right but doesn't spoil the intended point here because there are many mundane epistemic notions (e.g. well-grounded, supported by evidence etc.) that are ordinarily used in the sense of epistemic justification. These mundane expressions are equivalent to the more technical notion of epistemic justification. For this point see also R.Wedgwood (2009).

decision possibly will be one of life and death, as a false prescription might result in the death of the patient. Such a life-threatening case might be that of a patient with a harmless flu but with a record of serious allergy in antibiotics. This switch of epistemic habits (and epistemic norms) happens because our justification assertions and attributions are sensitive to the high or low practical stakes involved[184].

However, not only the same epistemic agent may endorse different epistemic habits in different contexts but also different epistemic agents might have different epistemic habits in the very same contexts. They might have different epistemic habits because they see the same context as posing different epistemic demands. This happens because different epistemic agents may have different epistemic goals. As William James (1979) famously pointed out, there are different epistemic goals an agent may pursue and this affects what one thinks he ought to believe. As James claimed, one may value avoiding error over attaining the truth while others might value attaining the truth over avoiding error.

These two epistemic goals seem to be in tension as the goal for attaining truth calls for more relaxed epistemic norms that would permit a number of false beliefs on the condition that we are to attain at least some true beliefs. Instead, the goal of avoiding error calls for more demanding epistemic norms that would not permit any false beliefs even if that would allow the attainment of a number of true beliefs. Accordingly, some agents in certain contexts may raise the demandingness of norms for justification because they might have the epistemic goal of avoiding error. Other agents in certain contexts may be more lax and lower the demandingness of their epistemic norms because they might have the epistemic goal of attaining truth rather than avoiding error.

Here is an example where agents have different epistemic habits in the same context because they entertain different epistemic goals. A religious person may be a coherentist with the epistemic habit to take any belief as justified if it coheres with his religious background web of beliefs. His religious beliefs are in the centre of his web of beliefs where the more certain beliefs (like those of logic and mathematics) lie. This agent thinks coherence with the rest of his religious beliefs is sufficient to license any belief as justified. Thus, for him the belief that 'angels and demons exist'

_____

[184] Such cases of 'pragmatic encroachment' are discussed in J.Stanley (2005) and E.Craig (1990).

is justified because he has the epistemic habit to consider anything licensed by the norm of coherence as justified. In this sense, he is an agent with the goal of attaining truth rather than avoiding error as he employs more relaxed and permissive epistemic norms.

Instead, a non-religious agent may have the epistemic habit to take any belief as justified only if it is licensed by the norm of epistemic duty where one's epistemic duty is to endorse only what is licensed by adequate empirical evidence. He thinks that religious beliefs are not based on adequate empirical evidence and therefore considers it as his duty to take 'the belief that angels and demons exist' as unjustified. It is unjustified because it is not licensed by the epistemic norm of empirical evidence fixed by the relevant epistemic habit.

In this sense he is an agent with the goal of avoiding error rather than attaining truth as he employs more demanding and restrictive epistemic norms. Thus, the employment of epistemic habits is not only sensitive to the different epistemic demands different epistemic contexts might pose but it is also sensitive to the kind of epistemic goals the agent has and the way these goals might imply more lax or more stringent epistemic norms for justification[185].

In addition, not only different epistemic habits may be endorsed in different contexts but also different *kinds* of habits may be endorsed in different contexts by epistemic justification assertions and attributions. One obvious species of habits can be distinguished on the basis of *the strength of the psychological grip of the habit*. In some cases the habits endorsed might be *deeply-rooted*, the product of rigorous internalization during childhood (which as everyone knows is a very malleable period). Such deeply-rooted habits will have a very strong grip on the agent and will be very difficult to get rid of them.

Some of these deeply-rooted habits maybe consciously endorsed by the agent, if the agent is sufficiently self-reflective and of a philosophical bent. The agent might acknowledge that he has such a habit and approve of that habit. But it is fair to assume that in most cases deeply-rooted habits are probably unconscious and the agent unbeknownst to him forms justified beliefs on their principle. This is probably the case because it is hard to keep track of one's epistemic habits, especially if one is

---

[185] For some discussion of this point see Craig (1990).

absorbed in the non-academic concerns of everyday life. In other cases the agent might acknowledge the existence of the habit and refuse to endorse or approve it but still be in the grip of the habit[186]. The habit might have grown so strong and deeply-rooted that the agent by now cannot control the exercise of this habit. Intuitively, deeply-rooted habits are very difficult to get rid of and, theoretically at least, such irrationality cases might occur.

Further, in yet other cases we might have *newly-established* habits. That is, a habit might have been recently established and consciously or unconsciously still be cultivated on behalf of the agent by practice. Reasonably, the grip of such habits on the psychology of the agent will be significantly weaker than the grip of the deeply-rooted habits. In cases of such newly-established habits there might be even instances where the agent consciously motivates and cultivates irrational habits for some pragmatic reason.

Think, for example, of an agent who realizes that he is afraid of heights because he is a very bad jumper. On the face of this realization, he might intentionally cultivate an epistemic habit of forming pragmatically justified beliefs that support the conviction that he is a great jumper. He might intentionally cultivate such an irrational epistemic habit because he foresees that belief in the improvement of his jumping ability will mitigate his heights phobia.

Also, there are *inchoate* habits that are just now starting to be consciously or unconsciously cultivated and therefore they might initially seem 'out of (habitual) character'. These habits will have a relatively weak grip on the psychology of the agent and will be the easiest to get rid of. Such inchoate habits surely exist and this might account for the intuition that in some cases we don't form justified beliefs on the basis of an existent epistemic habit or even form justified beliefs positively against an epistemic habit. In such a case what really happens is that we act or form justified beliefs on the basis of a new inchoate habit that has just started being cultivated and is yet a long way from being a typical deeply-rooted habit (or even a newly-established habit). In such cases there might be a 'struggle for ascendancy' between habits, as habits will struggle to gain the primary grip on the agent's psychology and finally utterly expel the rival habit.

---

[186] Such cases are reminiscent of the irrationality cases H.Frankfurt (1971) has described.

So far, on the basis of the strength of the psychological grip of the habit we distinguished three different kinds of habits that might endorse in justification discourse: deeply-rooted, newly-established and inchoate habits. Yet another interesting way of distinguishing a species of epistemic habits is on the basis *of amenability to learning*. On the basis of this property we may draw the distinction between *constitutive* and *acquired* epistemic habits[187].

Constitutive habits are ones that although 'triggered' by learning within the margins of a sociolinguistic community, they do not really seem amenable to learning. That is, although these habits lie dormant and are triggered to manifestation by learning (and other social stimulation), they don't seem to be revisable according to the dictates of various forms of learning. In this sense, once constitutive habits are being triggered and cultivated for some time they become deeply-rooted habits. They become habits that are extremely difficult, if not impossible to get rid of.

Such constitutive epistemic habits paradigmatically involve mathematical and logical justification (and perhaps perceptual justification too). Mathematical and logical justification seem to be paradigmatic examples of regulation by constitutive epistemic habits because almost any agent with a proper conceptual development in the social context of a linguistic community seems to have non-amenable-to-learning and broadly similar mathematical and logical norms[188].

On the one hand, mathematical and logical justified beliefs seem non-amenable to learning. For if we try to teach that, let us say, '1+1=3' and that 'a triangle is four-sided' this will be resisted even if the agent tries hard for the opposite. Even primary school children, I suppose, if inculcated to believe that '1+1=3' or that 'a triangle has four sides' at some point will resist these as counter-intuitive. They will count fingers and find them two, they will draw triangles and find them three-sided etc.

On the other hand, there is relative general agreement about logical and mathematical norms. Few agents would dare question, for example, that 'The belief that 1+1=2 is justified', that 'The belief that a square has four equal sides is justified' or that 'The belief that necessarily p and not-p is false is justified'. These simply strike us as deeply intuitive and, presumably, that is why foundationalists who appeal

---

[187] I borrow the distinction between constitutive and acquired habits from C.S.Peirce (1991).
[188] Feral children or other cases of agents who didn't have proper conceptual development in the social context of a linguistic community, put to one side.

to rational intuitions (e.g. L.Bonjour 1998) tend to appeal to maths and logic for convenient examples[189].

These features of constitutive epistemic habits might incline people to appeal to nativist ideas in order to account for their reality. For instance, someone could wield a Chomskian 'poverty of stimulus argument' for mathematical and logical justification. That is, as N.Chomsky (1980) famously argued that the richness and complexity of our implicit grammatical knowledge cannot be wholly explained by appeal to learning of grammatical rules because this is relatively impoverished, we could argue along the same lines for maths and logic. Our richness and complexity of mathematical and logical justification and our capacity to constantly generate new and diverse mathematical and logical justification cannot be wholly explained by appeal to learning and, therefore, we should postulate innate mathematical and logical faculties.

To be sure, there might be more species of habits to be distinguished but for present purposes this much of conceptual digging is I think enough. It is sufficient to make the intended point, namely, stress the complexity of epistemic habits and norms endorsed in epistemic justification discourse. In the next section we quickly consider two objections to habits-endorsment expressivism: 'the over-intellectualization objection' and 'the situationist challenge'.

7.6 Two Anticipated Objections and Quick Replies

In this final section, I would like to very quickly fend off two objections that might have arisen to the minds of some readers. No doubt, these objections are far from exhaustive and even for the two ones discussed much more could be said. But I think that the quick responses presented here will help stifle the initial appeal of these objections. I admit, though, that a more comprehensive and detailed response is needed and should come elsewhere for these objections.

The first one is 'the over-intellectualization objection'. The habits-endorsement expressivist picture of the semantics I have strived to sketch so far, might strike some people as hopelessly over-intellectualizing ordinary epistemic practice; a charge

---

[189] Of course, after turning an expert in mathematics or logic an agent might start scrutinizing the relevant mathematical and logical norms in play. This is presumably how non-Euclidean geometries and non-classical logics have been conceived. But scrutiny come at a level of high expertise and does not really affect the point being made here, namely, that these mathematical and logical habits are constitutive.

often made against many philosophical theories. For to claim that justification assertions and attributions express, -among other things like commendation that I didn't touch - this mouthful: endorsement of the habits of belief-fixation in virtue of which S believes that p where these habits of belief-fixation employ and are constrained by certain epistemic norms etc. etc. seems to over-intellectualize and therefore distort the everyday practice of justification assertions and attributions.

Yet, this objection is too quick. For no one wants to suggest that ordinary agents are really aware of the semantic implications of what they think and say. Indeed, few agents are really aware of their epistemic habits and norms. As Kant (2003) would have put it, this comes to the surface in the process of reason's journey of self-knowledge. And this journey of reason knowing itself, insofar it is relentlessly systematic, is an academic philosopher's job. I conclude therefore that this objection misses its mark. We are not in conflict with common epistemic practice, we rather relentlessly analyse common epistemic practice and arrive to such conclusions that seem too far from ordinary epistemic experience. But this does nothing to vitiate our proposal.

A second objection comes from the so-called situationist challenge against virtue theories. Recently, a number of philosophers have drawn empirical data from social psychology experiments in order to argue that the central virtue-theoretic notion of character is illusory (both for philosophers and laymen)[190]. These experiments, predominantly S.Milgram's (1974) famous obedience experiments, are taken to show that in reality the notion of character does not exist. It does not exist because in these experiments the agents' action is not guided by their character and their relevant virtues and traits but by the situational context they are found in.

These philosophers insist that although of course we have and use the concept of character in philosophical and everyday discourses, the concept has no referent and it is of the same order as unicorns and Santa Claus are. If true, it is obvious what dire implications this view has for approaches like habits-endorsement expressivism that rely on the virtue-theoretic notion of character. If true, such approaches are based on an illusion about human psychology.

However, things are not all that bleak for such approaches because the

---

[190] See for example G.Harman (1999).

situationist challenge seems to get things wrongly. Intuitively, these philosophers draw the wrong conclusion from the experimental data they adduce. The experimental data they cite simply shows at least some agents' actions to be guided by the situational context they are being found in. From this, they jump to the conclusion that character does not really exist, as what guides the conduct of agents is not character but simply the situational context they are found in.

Yet, these experimental results do not yield the conclusion that these philosophers are eager to draw, namely, that character is mere fiction and does not exist. First of all, if seen from the right angle these experimental results seem to corroborate rather than refute what virtue-theoretic approaches predict about human action and psychology. For virtue-theoretic approaches suggest that is very difficult for someone to have a really virtuous character and, therefore, the character inconsistencies found in these experiments are to be expected by virtue theory's own lights. Besides, that is why we think that being a virtuous agent is an important achievement to some important extent creditable to the agent himself. It is because few can come to be really virtuous.

Second, the notion of character is consistent with character inconsistencies. It is only the notion of virtuous character that is not. Many people have weak characters, as we say, but this doesn't mean that they don't have a character. This is exactly what we mean in ordinary discourse when we talk about weak characters or weak-willed persons. They are persons that do not seem to have a consistent behaviour, as their conduct seems to be fluctuating, prone to be drifted away by the current of the situational context they might find themselves in. Sometimes they are just other times unjust, sometimes patient other times impatient, sometimes self-restraint other unrestrained and so on. They might exhibit brave behaviour in some circumstances and cowardice in others. The might exhibit self-restraint in relation to chocolate but not to red wine etc.

Third, that the situationist challenge goes too far is obvious from the fact that character is an indispensable notion deeply interwoven in folk-psychology and reasons-explanations. Both in philosophical and ordinary discourse, in explaining intentional action or belief-fixation we do not merely appeal to reasons for action or reasons for belief. We also quite often appeal to the character of the agent in

question. If, for example, I know that Mike is a person who is committed to the pursuit of argument wherever it leads then I can predict that he will endorse the conclusion of a compelling argument. Again, if I know that Matthew is a person who highly values human life, then I can predict that he will save a boy from drowning in a shallow pond despite the muddied clothes.

This brief and, admittedly, inadequate to do full justice to these (and other) objections section marks the end of the discussion. Let us, finally, turn to the conclusion and summary of the argument.

7.7 Conclusion and Summary of the Argument

I have argued for a novel version of norm-expressivism dubbed 'habits-endorsement expressivism'. The starting point of my argumentation for habits-endorsement expressivism was Gibbard's (2003, 2008) theory of plan-reliance expressivism for knowledge discourse. I introduced the theory, extended it to cover the epistemic justifcation discourse and then identified some of its problems. As I have argued, plan-reliance expressivism cannot account for the nonvoluntarism intuition while it also runs into direct counterexamples.

As a response to these discontents, I introduced habits-endorsement expressivism by means of analysing what we seem to implicate in discursive contexts of justified belief attribution. Afterwards, I went on to argue that habits-endorsement is more explanatory powerful from plan-reliance expressivism because it can keep the attractions of both norm-expressivism and plan-reliance expressivism while at the same time outflank the identified problems of plan-reliance expressivism. Next, I added some more touches on my initial sketch of habits-endorsement expressivism and finally anticipated two quick objections to habits-endorsement expressivism and tried to provide two quick responses meant to stifle the initial appeal of these objections. These objections were 'the over-intellectualization objection' and 'the situationist challenge'.

Yet, this is not the end of the road. For even if my version of expressivism evades these objections, there are still other important semantic challenges (the Frege-Geach problem, the truth problem etc.) to face that go straight to the heart of the semantic programme of expressivism. These semantic challenges suggest that expressivist theories cannot be considered as full-blown semantic theories unless they meet these

challenges successfully. If they don't manage to meet these semantic challenges, then all expressivist theories eventuallly amount to are pragmatic theories. That is, theories of the mental content expressed in various contexts by speakers but not of linguistic content per se.

I don't have much time to discuss these notorious semantic challenges at any depth, but as neglecting these important challenges to expressivism might give the wrong impression to readers, namely, that I don't take the challenges seriously enough or, even worse, that I ignore the challenges the Epilogue is a quick note to these semantic challenges.

Epilogue The Semantic Problems of Normative Expressivism: A Short Note

In the last chapter 7 we went some way to argue for a habits-endorsement expressivist solution to 'the epistemic justification puzzle'. We argued for this solution by means of claiming that it can both capture the attractions of norm-expressivism and plan-reliance expressivism while outflank the problems of plan-reliance expressivism.

But as it is common wisdom, in philosophy, rarely if ever philosophical explorations have an all-perfect ending. As we have already hinted in the concluding paragraphs of the Ch.7, normative expressivist approaches face semantic challenges that threaten to strangle the expressivist semantic programme from its infancy. That is, semantic challenges that if expressivism is to establish itself as a full-blown semantic theory, these challenges should be addressed. This last short chapter is a note to these semantic challenges. Let us turn to this.

Semantic orthodoxy has it that semantic theories are referential and truth-conditional. They explain the meaning of sentences in terms of linguistic content and what that purports to refer to. And they are referential and truth-conditional for a very good reason. The reason is that truth-conditional semantics can easily account for the logical properties of complex sentences embedded in unasserted truth-functional contexts, as the meaning of sentences is 'fixed' by the truth-conditions of the sentence. The meaning of the sentences is determined by their truth-conditions and truth-conditions do not allow any shift of meaning that would render a valid inference invalid due to a fallacy of equivocation. That is, a fallacy where an invalid inference is drawn on the basis of using the same words with different meanings from one premise to another.

Here is an example of a fallacious modus ponens inference because of equivocation.

P1: 'Feathers are light'

P2: 'If something is light, then it is not dark'

C: 'Feathers are not dark'.

This equivocation fallacy plays on the two different meanings of 'light' employed

in P1 and P2, namely, 'light' in the sense of not heavy and 'light' in the sense of having bright colour. Truth-conditional semantics rule out such occurrences because they do not allow for any such shift of meaning from premise to premise. As they 'fix' the meaning of sentences with truth-conditions they rule out the possibility of any shift of meaning that would have led to a fallacy of equivocation.

For example, take this modus ponens:

P1:'Heliocentricism is justified'

P2 : 'If heliocentricism is justified, then you should instruct your little brother to believe it'

C: 'You should instruct your little brother to believe it'

This is a valid inference and truth-conditional semantics can easily explain why. Truth-conditions fix the meaning of the sentences and this guarantees that the meaning of P1 will remain the same in the antecedent of P2 and the conclusion will follow. Sameness of meaning guarantees truth-preservation and truth-preservation forces the valid conclusion. That is, if P1 is true and P2 is true, then C *must* be true[191]. And all this happens in a thoroughly de-psychologized way, as all hinges on the impersonal linguistic content and nothing on the personal mental content of a speaker.

However, in contrast to truth-conditional theories of semantics, expressivist theories are psychologistic theories of semantics. They are theories that analyse the meaning of sentences not in terms of linguistic content and reference/truth-conditions but in terms of the mental content expressed by speakers. This seems to make expressivism prima facie inconsistent with a theory of semantics, which, as Frege (1997) insisted, it has to be de-pshychologized. For, semantics ought to be de-psychologised and expressivism is in essence a psychologistic theory of meaning.

This is the real point behind P.Geach's (1972) notorious *Frege point*, now very well known as 'the Frege-Geach Problem' (or 'the embedding problem'). The Frege-Geach problem suggests that expressivist theories cannot establish themselves as full-blown semantic theories because they can't account for the logical properties of complex sentences embedded in unasserted truth-functional contexts, as any theory

---

[191] As L.Caroll (2008) has made clear in his famous 'What the Turtoise said to Achilles' we can start an infinite regress for justifying our valid inferences. I take the standard view here and ignore such ' infinite regress' complications.

of semantics ought to do. Expressivism cannot account for the logical properties of normative discourse because, as a psychologistic theory of meaning, it allows for shift of meaning from sentence to sentence and this means that valid conclusions won't follow in virtue of a fallacy of equivocation. If that is the case, then our analysis of justification assertions and attributions is not adequate for a natural language where assertions and attributions can be found embedded in logically complex truth-functional contexts. Let me elaborate.

Expressivism explains meaning in terms of mental content (or 'states of mind'). To say that 'murder is wrong' or that 'p is justified' is to express a state of mind. What exactly is this state of mind depends on the contours of the expressivist theory one has. Yet, we also use normative language in truth-functional, unasserted contexts like in conditionals, disjunctions, negations etc. The problem, however, is that in such 'unasserted contexts' we might not express the noncognitive attitude we express in contexts of simple assertion. This suggests that in such cases we are using the sentences with different meaning from one premise to another and this thwarts sameness of meaning that is necessary for validity. The upshot is that valid inferences are rendered invalid.

For example, take this epistemic disjunctive syllogism:

P1: 'Geocentricism is not justified.'

P2: 'Either heliocentricism is justified or geocentricism is justified'.

C: 'Therefore, heliocentricism is justified.'

This is a valid argument, if anything is, and the problem is that if we assume expressivism as our semantic theory then it seems that there will be cases where the valid conclusion won't follow because the speaker might express different meaning in P1 from the second disjunct in P2. P1 is an assertion and according to, say, norm-expressivism, we express an attitude of endorsement for the epistemic norms that rule out this belief. P2, though, it is not an assertion and in this case it seems wide open that a speaker might utter the sentence without expressing any attitude of norm-endorsement for the meaning of the second conjunct. As Frege (1997) would have said, we can utter P2 without expressing any 'assertoric force'. In such an event, we would have a fallacy of equivocation.

Now, if expressivism cannot provide a response to the Frege-Geach problem and

provide semantics for natural languages that can account for validity (and overall logical properties) then expressivism is doomed. It is doomed because a theory that can account only for atomic normative sentences like assertions and attributions but not for logically complex sentences cannot model the workings of natural languages (which, obviously, are logically complex). For, any semantic theory must be able to account for things like 'if p then q; p; therefore q' 'p; therefore $\neg\neg p$' and an infinite number of other logically complex inferences.

In the face of the Frege point, expressivists have tried a number of ways to resolve the Frege-Geach problem. Some have tried to build a 'logic of attitudes', others have incorporated a belief component into the mental state expressed meant to do the necessary semantic work etc[192]. Reasonably, we cannot even begin to introduce these approaches here or moreover to examine whether they work or not. The thesis has proceeded on the hopeful assumption that somehow these problems can be resolved and that, if such a solution is discovered, habits-endorsement expressivism can follow such a solution.

But even with solving the Frege-Geach problem this won't mean that expressivists are done with their problems with logic. This is an important point M.Schroeder (2008) has drawn. M.Schroeder (2008) has argued himself for a certain solution to the Frege-Geach problem but also pointed out that for expressivism to establish itself as a full-blown semantic theory is not enough to just solve the Frege-Geach problem. The reason is that normative sentences do not just appear embedded in unasserted truth-functional contexts with other normative sentences. They also appear in such contexts with ordinary descriptive sentences like 'grass is green'. We can have, for instance, the sentence ' Either grass is green or Heliocentricism is not justified'.

This means that expressivists need to explain the workings of natural languages in all sorts of discourse and not just the normative discourses. They need to explain how language ought to work in its entire field of application, namely, normative, ordinary descriptive and their combination. This means that for normative expressivism to

---

[192] Logic of attitudes approaches have been suggested by S.Blackburn (1984, 1993, 1998) and A.Gibbard (1990, 2003). M.Ridge (2006) has proposed that expressivim can solve the problem if it goes 'ecumenical', that is, incorporate a belief into the content of what is being expressed. M.Chrisman (2010) has argued that inferentialism could solve the problem and that this makes it superior to expressivism.

work it must expand to cover the workings of natural language in general. It must offer an expressivist picture of language *tout court*. This is a big task and if expressivism can accomplish such a task, then, indeed, we can a la Kant talk of expressivism as effecting a 'Copernican revolution' in philosophy. But whether expressivism can meet this semantic challenge is an open question for conceptual exploration. Our own journey of conceptual exploration, however, ends here.

References

Alston, W. (2005). *Beyond Epistemic Justification*. Ithaca, Cornell University Press.

Anscombe, G.E.M. (2000). *Intention*. Cambridge MA, Harvard University Press.

Aristotle. (2003a). *Nicomachean Ethics*. Cambridge MA, Harvard University Press.

----------. (2003b). *Metaphysics Books I-IX*. Cambridge MA, Harvard University Press.

----------. (2006). *Metaphysics Books X-XIV*. Cambridge MA, Harvard University Press.

Armstrong, D. M. (1980). *A Theory of Universals*, Vol.I-II. Cambridge, Cambridge University Press.

Audi, R. (2001). 'Doxastic Voluntarism and the Ethics of Belief' in *Knowledge and Truth and Duty*, (ed.) M.Steup. Oxford, Oxford University Press. pp. 93-111.

----------. (2007). 'Intuitionism, Pluralism and the Foundations of Ethics' in *Foundations of Ethics,* (eds) R.Shafer-Landau and T.Cuneo. Oxford, Blackwell Publishing. pp. 402-420.

Austin, J. (1961). 'Other Minds' in his *Philosophical Papers*, (eds.) J. O. Urmson and G. Warnock. Oxford, Oxford University Press. pp. 76-117.

-----------. (1962). *How to Do Things With Words*. Oxford, Oxford University Press.

Avraamides, A. (1999). 'Intention and Convention' in *A Companion to the Philosophy of Language*, (eds.) B.Hale and C.Wright. Oxford, Blackwell Publishing. pp. 60-86.

Ayer, A. J. (1946). *Language Truth and Logic*. London, Penguin Books.

Baker, J. (2007). *50 Physics Ideas*. London, Quercus Publishing.

Baldwin, T. (2006). 'The Three Phases of Intuitionism' in *Ethical Intuitionism*, (ed). P. Stratton-Lake. Oxford, Oxford Univerity Press. pp. 92-113.

Bealer, G. (1998). 'Intuition and the Autonomy of Philosophy' in *Rethinking Intuition*, (eds.) M. DePaul and W.Ramsey. London, Rowman and Littlefield Publishers. pp. 201-239.

Bennett, K. and MacLaughlin, B. (2005). 'Supervenience' in *The Stanford Encyclopedia of Philosophy*, (ed.) E.N.Zalta.
URL: http://plato.stanford.edu/entries/supervenience/

Blackburn, S. (1984). *Spreading the Word*. Oxford, Oxford University Press.

---------------. (1993). 'Morals and Modals'. pp. 52-75.

---------------. (1993). 'Knowledge, Truth and Reliability'. pp. 35-52.

---------------. (1993). 'Moral Realism'. pp. 111-130.

---------------. (1993). 'Supervenience Revisited'. pp. 130-149.

---------------. (1993). 'Attitudes and Contents'. pp. 182-198.

---------------. (1993). 'Errors and the Phenomenology of Value'. pp. 149-166.

---------------. (1993). 'How to be an Ethical Antirealist'. pp. 166-182.

All in his *Essays in Quasi-Realism*. (1993). Oxford, Oxford University Press.

---------------. (1996). 'Securing the Nots: Moral Epistemology for the Quasi-Realist' in *Moral Knowledge?,* (eds.) W. Sinnott-Armstrong and M.Timmons. Oxford, Oxford University Press. pp. 82-100.

---------------. (1998). *Ruling Passions*. Oxford, Oxford University Press.

Bonjour, L. (1985). *The Structure of Empirical Knowledge*. Cambridge, MA Harvard University Press.

-------------. (1998). *In Defense of Pure Reason*. London, Cambridge University Press.

Boyd, R. (2007). 'How to be A Moral Realist' in *Foundations of Ethics*, (eds.) R.

Shafer-Landau and T.Cuneo. Oxford, Blackwell Publishing. pp. 163-186.

Brandom, R. (1995). 'Knowledge and the Social Articulation of the Space of Reasons'. *Philosophy and Phenomenological Research* LV (No. 4). pp. 895-908.

----------------. (2000). *Articulating Reasons*. Cambridge MA, Harvard University Press.

Brink, D. (1989). *Moral Realism and the Foundations of Ethics*. Cambridge, Cambridge University Press.

Caroll, L. (2008). 'What the Turtoise Said to Achilles' in *The Complete Illustrated Lewis Caroll*. London, Wordsworth. pp. 1178-1182.

Carter, A. and Chrisman, M. (2010). 'Is Expressivism Incompatible with Inquiry?'. Unpublished Manuscript.

Chalmers, D. (1996). *The Conscious Mind*. Oxford, Oxford University Press.

Chisholm, R. (1966). *Theory of Knowledge*. Englewood Cliffs, Prentice Hall.

Chomsky, N. (1980). *Rules and Representations*. Oxford, Blackwell Publishing.

Chrisman, M. (2007). 'From Epistemic Contextualism to Epistemic Expressivism'. *Philosophical Studies* Vol.135(2).

----------------. (2008). 'Expressivism, Inferentialism and Saving the Debate' *Philosophy and Phenomenological Research* Vol. 77(2).

------------------. (2010). 'From Epistemic Expressivism to Epistemic Inferentialism'. *Social Epistemology*. A. Millar, D. Pritchard and A.Haddock. Oxford, Oxford University Press.

Churchland, P. (2003). 'Eliminative Materialism and the Propositional Attitudes' in *Mind and Cognition*, (ed.) W.Lycan. Oxford, Blackwell Publishing. pp. 120-133.

Conee, E. and Feldman, R. (2004). 'Evidentialism' in their *Evidentialism.* Oxford, Oxford University Press. pp. 83-111.

Craig, E. (1990). *Knowledge and the State of Nature*. Oxford, Oxford University Press.

Crane, T. (2001). *Elements of Mind*. Oxford, Oxford University Press.

Cuneo, T. (2007). *The Normative Web*. Oxford, Oxford University Press.

Dancy, R. (2004). *Plato's Introduction of the Theory of Forms*. Cambridge, Cambridge University Press.

Davidson, D. (1996). 'Truth and Meaning' in *The Philosophy of Language*, (ed.) A.P.Martinich. Oxford, Oxford University Press. pp. 92-104.

----------------. (2001a). 'The Folly of Trying to Define Truth' in *The Nature of Truth*, (ed.) M.Lynch. Hong Kong, MIT. pp. 623-641.

----------------. (2001b). 'Actions Reasons and Causes' in his *Essays on Actions and Events*. Oxford University Press. pp. 3-21.

----------------. (2003). 'Mental Events' in *Mind and Cognition,* (ed.) W.Lycan. Oxford, Blackwell Publishing. pp. 35-46.

----------------. (2008). 'A Coherence Theory of Truth and Knowledge' in *Epistemology*, (eds.) E.Sosa, J.Kim, J.Fantl and M.McGrath. Oxford, Blackwell Publishing. pp. 124-134.

Davis, W. (2010) 'Implicature' in *The Stanford Encyclopedia of Philosophy,* (ed.) E.N. Zalta . URL : http://plato.stanford.edu/entries/implicature/

Descartes, R. (2008). *Meditations on First Philosophy*. Oxford, Oxford University Press. Translated with an introduction by M.Moriarty.

Dennett D. (2003). 'True Believers: Intentional Strategy and Why it Works' in *Mind and Cognition,* (ed.) W.Lycan. Oxford, Blackwell Publishing. pp. 75-87.

De Vries W. (2005). *Wilfrid Sellars*. Chesham, McGill-Queen's University Press.

Dostoyevsky, F. (2003). *The Brothers Karamazov*. London, Penguin Books. Translated with an introduction by D.McDuff.

--------------------. (2008). *Memoirs from the House of the Dead*. Oxford, Oxford University Press. Translated by J.Coulson.

Dreier, J. (2004). 'Meta-Ethics and the Problem of Creeping Minimalism'. *Philosophical Perspectives* 18. pp.23-44.

Evans, N. and Levinson, S. C. (2009). 'The Myth of Language Universals: Language Diversity and its Importance for Cognitive Science'. *Behavioral and Brain Sciences* 32(5). pp. 429-492.

Feldman, R. (2001a). 'Naturalized Epistemology' in *The Stanford Encyclopedia of Philosophy*, (ed.) E.N.Zalta.
URL = http://plato.stanford.edu/entries/epistemology-naturalized/

----------------. (2001b). 'Voluntary Belief and Epistemic Evaluation' in *Knowledge, Truth and Duty*, (ed.) M.Steup. Oxford, Oxford University Press. pp. 77-93.

Field, H. (2000). 'Apriority as an Evaluative Notion' in *New Essays on the A Priori*. (eds.) P. Boghossian. and C.Peacocke. Oxford, Oxford University Press. pp. 117-150.

----------. (2009). 'Epistemology without Metaphysics'. *Philosophical Studies* Vol. 143(2). pp. 249 -290.

Firth R. (2007). 'Ethical Absolutism and the Ideal Observer' in *Foundations of Ethics*, (eds.) R.Shafer-Landau and T.Cuneo. Oxford, Blackwell Publishing. pp. 107-120.

Fodor, J. (1975). *The Language of Thought*. Cambridge, Cambridge University Press.

----------. (1998). 'Review of J.McDowell's Mind and World' in *In Critical Condition*. Cambridge, The MIT Press. pp. 3-9.

Fogelin, R. (1994). *Pyrrhonian Reflections on Knowledge and Justification*. Oxford, Oxford University Press.

Foley, R. (2002). 'Conceptual Diversity in Epistemology' in *The Oxford Handbook to Epistemology*, (ed.) P.Moser. Oxford, Oxford University Press. pp. 177-204.

Frankena, W. (1939). 'The Naturalistic Fallacy'. *Mind* 48. pp. 464-477.

Frankfurt H. (1971). 'Freedom of the Will and the Concept of a Person'. *Journal of Philosophy*. Vol.68(1). pp. 5-20.

Frege, G. (1997). 'Thought'. pp. 325-346.

-----------. (1997). 'On Sinn and Bedeutung'. pp. 151-177.

-----------. (1997). 'Negation'. pp. 346-342.

All in *The Frege Reader*, (ed.) M.Beaney. Oxford, Blackwell Publishing.

Fumerton, R. (2002). 'Theories of Epistemic Justification' in *The Oxford Handbook to Epistemology*, (ed.) P.Moser. Oxford, Oxford University Press. pp. 204-234.

----------------. (2010). 'Foundationalist Theories of Epistemic Justification' in *The Stanford Encyclopedia of Philosophy*, (ed.) E.N.Zalta.
URL = http://plato.stanford.edu/entries/justep-coherence/

Geach, P. (1972). 'Assertion' in his *Logic Matters*. Oxford, Basil Blackwell Publishing. pp. 250-274.

Gettier, E. (1963). 'Is Justified True Belief Knowledge?'. *Analysis* 23. pp. 121-3.

Gibbard, A. (1990). *Wise Choices Apt Feelings*. Oxford, Oxford University Press.

---------------. (2003). *Thinking How to Live*. Cambridge, MA, Harvard University

Press.

--------------. (2008). *Reconciling Our Aims*. Oxford, Oxford University Press.

Ginet, C. (2001). 'Deciding to Believe' in *Knowledge Truth and Duty*, (ed.) M.Steup. Oxford, Oxford University Press. pp. 63-77.

Goldie, P. (2002). *The Emotions*. Oxford, Oxford University Press.

Goldman, A. (1978). 'Epistemology and the Psychology of Belief'. *Monist* 61. pp. 525-535.

---------------. (1986). *Epistemology and Cognition*. Cambridge MA, Harvard University Press.

--------------. (1991). 'What is Justified belief?'. pp. 105-127.

--------------. (1991). 'Strong and Weak Justification'. pp.127-143.

Both in his *Liaisons*. Cambridge, MA, The MIT Press.

---------------. (2002). 'A Priori Warrant and Naturalistic Epistemology' in his *Pathways to Knowledge*. Oxford, Oxford Univresity Press. pp. 24-51.

Greco J. (2003). 'Knowledge as Credit for True Belief' in *Intellectual Virtue: Perpectives from Ethics and Epistemology*, (eds.) M.DePaul and L.Zagzebski. Oxford, Oxford University Press.

Grice, P. (1989). 'Meaning'. pp. 213-224.

----------. (1989). 'Logic and Conversation'. pp. 22-41.

----------. (1989). 'Further Notes on Logic and Conversation'. pp. 41-57.

----------. (1989). 'Utterer's Meaning and Intentions'. pp. 86-117.

----------. (1989). 'Prolegomena'. pp . 3-22.

All in his *Studies in the Ways of Words*. (1989). Cambridge MA, Harvard University Press.

Hajek A. (2004). 'Pascal's Wager' in *The Stanford Encyclopedia of Philosophy*, (ed.) E.N. Zalta. URL = http://plato.stanford.edu/entries/pascal-wager/

Hare, R. M. (1952). *The Language of Morals*. Oxford, Oxford University Press.

-------------. (1963). *Freedom and Reason*. Oxford, Oxford University Press.

Harman, G. (1986). *Change in View*. Cambridge MA, The MIT Press.

-------------. (1999). 'Moral Theory Meets Social Psychology: Virtue Ethics and the Fundamental Attribution Error'. *Proceedings of the Aristotelian Society Vol.* 99. pp. 315-31.

-------------. (2007). 'Moral Relativism Defended'. *Foundations of Ethics*, (eds.) R. Shafer-Landau and T.Cuneo. Oxford, Oxford University Press. pp. 84-93.

Helm, P. (1994). *Belief Policies*. Cambridge, Cambridge University Press.

Horgan, T. and.Timmons, M. (2007). 'New Wave Moral Realism Meets Moral Twin Earth'. *Foundations of Ethics*, (eds.) R. Shafer-Landau and T.Cuneo. Oxford, Blackwell Publishing. pp. 495-504.

Horwich P. (1990). *Truth* . Oxford, Basil Blackwell.

Howson, C. (1997). *Logic with Trees*, Oxford, Oxford University Press.

Hume, D. (1986). *A Treatise of Human Nature*. London, Penguin Books.

------------. (2005). *Enquiries*. Oxford, Oxford University Press.

Jackson, F. (1998). *From Metaphysics to Ethics*. Oxford, Oxford University Press.

James, W. (1979). 'The Will to Believe' in *The Works of William James*. Cambridge MA, Harvard University Press. pp. 13-34.

Joyce, R. (2007). 'The Myth of Morality' in *Foundations of Ethics*, (eds.) R. Shafer-Landau and T.Cuneo. Oxford, Blackwell Publishing. pp. 23-35.

Kahneman, D and Tversky, A. (1985). 'Judgments of and by Representativeness', in

Judgment Under Uncertainty: Heuristics and Biases, D.Kahneman, P.Slovic, and A.Tversky, (eds.). pp. 84-98.

Kant, I. (2003). *Critique of Pure Reason*. NewYork, Palgrave Macmilan.

--------. (2006). *Groundwork of the Metaphysics of Morals*. Cambridge, Cambridge University Press. With an introduction by C.Korsgaard.

Kappel, K. (2010). 'Expressivism About Knowledge'. Unpublished Manuscript.

-------------. (2010). 'Norm-Expressivism about Knowledge and the Value of Knowledge'. *Acta Analytica* Vol. 25(2). pp. 175-194.

Kearns, K. (2000). *Semantics*. Hampshire, Palgrave MacMillan.

Kim, J. (2000). *Mind in a Physical World.* Cambridge MA, The MIT Press.

--------. (2005). *Physicalism or Something Near Enough.* Princeton, Princeton University Press.

--------. (2008). 'What is 'Naturalized Epistemology'?'. *Epistemology*, (eds.) J. Kvanvig, E.Sosa, J.Fantl and M.McGrath. Oxford, Blackwell Publishing. pp. 538-552.

Klagge, J. (1988). 'Supervenience: Ontological and Ascriptive'. *Australasian Journal of Philosophy* Vol. 66(4). pp. 461-470.

Klein, P. (2008). 'How A Pyrrhonian Skeptic Might Respond to Academic Scepticism' in *Epistemology*, (eds.) E.Sosa, J.Kim, J.Fantl and M.MacGrath. pp. 35-51.

Kornblith, H. (2005). *Knowledge and its Place in Nature*. Oxford, Oxford University Press.

---------------. (2007). 'The Metaphysical Status of Knowledge' in *The Metaphysics of Knowledge,* (eds.) E.Sosa and E.Villaneuva. *Philosophical Issues* 17. pp. 145-165.

Korsgaard, C. (1996). *The Sources of Normativity*. Cambridge, Cambridge University Press.

Kripke, S. (1981). *Naming and Necessity*. Oxford, Blackwell Publishing.

Kuhn, T. (1996). *The Structure of Scientific Revolutions*. Chicago, University of Chicago Press.

Kvanvig, J. (2003). *The Value of Knowledge and the Pursuit of Understanding*. Cambridge, Cambridge University Press.

-------------. (2003). 'Coherentist Theories of Epistemic Justification' in *The Stanford Encyclopedia of Philosophy*, (ed.) E.N.Zalta.
URL = http://plato.stanford.edu/entries/justep-coherence/

Kyriacou, C. (2008). 'Naturalism and Normativity', (eds.) T.Achourioti, E.Andrade & M.Staudacher. Online Proceedings of the Graduate Conference on Normativity, Universiteit Van Amsterdam, The Netherlands. pp. 105-127 .
URL= http://www.illc.uva.nl/Publications/ResearchReports/X-2009-02.text.pdf

-------------. (2009). 'Plato, Moore and 'Ethical Nonnaturalism'', (ed.) P.Hanna. *An Anthology of Philosophical Studies*. Athens, ATINER. pp. 289-299.

-------------. (2009). 'Evolutionary Ruminations on 'the Value of Knowledge Intuition''. Fortcoming in the Proceedings of the 'Knowledge Value Evolution Conference', Prague, Nov. 2009.

-------------. (2010). 'Causal Exclusion and R.Shafer-Landau's Nonnaturalism'. Unpublished Manuscript.

-------------. (2010). 'Early Plato on Knowledge'. Unpublished Manuscript.

Lawrence, S and Margolis, E. (2003). 'Concepts and Conceptual Analysis'. *Philosophy and Phenomenological Research*. Vol. LXVII(2). pp. 253-282.

Lenman, J. (2003). 'Moral Deviants and Amoral Saints: A Dilemma for Moral Externalism'. *Southern Journal of Philosophy* 61. pp. 223-240.

Lewis, D. (1986). *The Plurality of Worlds*. Oxford, Blackwell Publishing.

-----------. (2004). 'New Work for A Theory of Universals' in *Metaphysics*, (eds.) J. Kim and E.Sosa. Oxford, Blackwell Publishing. pp.209-223.

Locke, J. (1973). *An Essay Concerning Human Understanding*. London, Collins Clear-Type Press.

Lycan, W. (2006). *Philosophy of Language*. New York, Routledge.

Lynch, M. (2009). 'Truth, Value and Epistemic Expressivism'. *Philosophy and Phenomenological Research* Vol. 79(1). pp. 76-97..

Mackie, J. (1977). *Ethics*. London, Penguin Books.

McDowell, J. (1994). *Mind and World*. Cambridge, Harvard University Press.

----------------. (1995). 'Knowledge and the Internal'. *Philosophy and Phenomenological Research* Vol. 55(4). pp. 877-893.

----------------. (1998). 'Virtue and Reason'. pp. 50-77.

----------------. (1998). 'Values and Secondary Qualities'. pp. 131-151.

----------------. (1998). 'Two Sorts of Naturalism'. pp. 167-198.

in his *Mind Value and Reality*. (1998). Cambridge MA, Harvard University Press.

-----------------. (2002). 'Knowledge and the Internal Revisited'. *Philosophy and Phenomenological Research* Vol. 64(1). pp. 97-105.

McGinn, C. (1999). *Ethics Evil and Fiction*. Oxford, Oxford University Press.

Milgram, S. (1974). *Obedience to Authority*. Hapercollins.

Mill, J. S. (1991). 'Utilitarianism'. *On Liberty and Other Essays*. Oxford, Oxford University Press. pp. 131-176.

Miller, A. (2005). *An Introduction to Contemporary Metaethics*. Cambridge, Polity Press.

-----------. (2006). *Philosophy of Language*. Oxford, Routledge.

Millikan, R. (2003). 'Biosemantics'in *Mind and Cognition*, (ed.) W.Lycan. Oxford, Blackwell Publishing. pp. 221-230.

Moore, G. E. (2000). *Principia Ethica*. Cambridge, Cambridge University Press. Edited with an introduction by T.Baldwin.

Moreland, J. P. (2001). *Universals*. Chesham, Acumen Publishing.

Mougin, A and Sober, E. (1994). 'Betting Against Pascal's Wager'. *Nous* Vol.28(3). pp. 382-395.

O'Shea J. (2008). 'Naturalism and Normativity : Sellars' 'Janus-Faced' Space of Reasons', (eds.) T.Achourioti, E.Andrade & M.Staudacher. Online Proceedings of the Graduate Conference on Normativity, Universiteit Van Amsterdam, The Netherlands. pp. 1-15. .
URL= http://www.illc.uva.nl/Publications/ResearchReports/X-2009-02.text.pdf

Owens, D. (2003). *Reason Without Freedom*. London, Routledge.

Papineau, D. (2003). 'Normativity and Judgement'. pp. 6-39.

----------------. (2003). 'The Evolution of Knowledge'. pp. 39-83.

Both in his *The Roots of Reason*. Oxford, Oxford University Press.

----------------. (2009). 'Naturalism'. *The Stanford Encyclopedia of Philosophy, (ed.) E.N.Zalta. URL =* http://plato.stanford.edu/entries/naturalism/

Pappas, G. (2005). 'Internalist Vs Externalist Conceptions of Epistemic justification' in *The Stanford Encuclopedia of Philosophy,* (ed.) E.N.Zalta.
URL = http://plato.stanford.edu/entries/epistemology/

Peacocke, C. (2004). *The Realm of Reason*. Oxford, Oxford University Press.

Peirce, C. S. (1991). 'The Fixation of Belief' in *Peirce on Signs*, (ed.) J.Hoopes. Chapel Hill, The University of North Carolina Press. pp. 144-160.

Plato (1928). *Theaetetus Sophist*. London, William Heinemann.

------. (2001). *Lysis Symposium Gorgias*. Cambridge MA, Harvard University Press.

------. (2003). *The Republic*. London, Penguin Books.

------. (2005). *Euthyphro Apology Crito Phaedo Phaedrus*. Cambridge MA, Harvard University Press.

Pollock, J. (1985). *Contemporary Theories of Knowledge*. London, Rowman and Littlefield.

Pritchard, D. (2007) 'The Value of Knowledge'. *The Stanford Encyclopedia of Philosophy, (ed*.) E.N.Zalta.
URL = http://plato.stanford.edu/entries/knowledge-value/

--------------. (2010). 'The Value Problem for Knowledge' in *The Nature and Value of Knowledge: Three Investigations,* co-authored with A.Haddock and A.Millar. Oxford, Oxford University Press.

Putnam, H. (1975). 'The Meaning of Meaning' in his *Mind Language and Reality*. Cambridge, Cambridge University Press. pp. 215-271.

-------------. (1978). *Meaning and the Moral Sciences*. London, Routledge and Kegan Paul.

-------------. (1981). *Reason Truth and History*. Cambridge, Cambridge University Press.

-------------. (2003). 'The Nature of Mental States' in *Mind and Cognition,* (ed.) W.Lycan. Oxford, Blackwell Publishing. pp. 27-35.

-----------. (2004). 'On Properties' in *Metaphysics*, (eds.) J. Kim and E.Sosa. Oxford, Blackwell Publishing. pp. 243-253.

Quine, W. V. (1953). 'Two Dogmas of Empiricism' in his *From a Logical Point of View*. Cambridge MA, Harvard University Press. pp. 20-47.

----------------. (2008). ' Naturalized Epistemology' in *Epistemology,* (eds.) E.Sosa, J.Kim, J.Fantl and M.MacGrath. pp. 528-538.

Rawls, J. (1999). *A Theory of Justice*, Harvard University Press.

Ridge, M. (2003) 'Moral Nonnaturalism' in *The Stanford Encyclopedia of Philosophy,* (ed.) E.N.Zalta.
URL= http://plato.stanford.edu/entries/moral-non-naturalism/

Ridge, M. (2006). 'Ecumenical Expressivism : Finessing Frege'. *Ethics* 116(22). pp. 302-336.

-----------. (2007a). 'Ecumenical Expressivism : The Best of both Worlds'. *Oxford Studies in Metaethics,* Vol.2 . R.Shafer-Landau. Oxford, Oxford University Press. pp. 302-336.

-----------. (2007b). 'Epistemology for Ecumenical Expressivists'. *Proceedings of the Aristotelian Society* Vpl. 81(1). pp. 83-108.

-----------. (2007c). 'Anti-Reductionism and Supervenience'. *Journal of Moral Philosophy* 4(3). pp. 330-348.

Rorty, R. (1997). *Philosophy and the Mirror of Nature*. Princeton, Princeton University Press.

Russel, B. (1956). 'On Denoting' in his *Logic and Knowledge*, (ed.) R. C. Marsh. London, George Allen and Unwin. pp. 39-57.

-----------. (2001). *The Problems of Philosophy*. Oxford, Oxford University Press.

Ryle, G. (1949). *The Concept of Mind*. Middlesex, Penguin Books.

Sayre-McCord, G. (1997a). ''Good' on Moral Twin Earth'. *Philosophical Issues* 8: 267-292.

----------------------. (1997b). 'The Metaethical Problem' *Ethics 108*. pp. 55-83.

Schroeder, M. (2008). *Being For*. Oxford, Oxford University Press.

----------------- . (2009). 'Hybrid Expressivism: Virtues and Vices'. *Ethics* 119(2). pp. 257-309.

Searle, J. (1958). 'Proper Names' *Mind* 67. pp. 166-173.

----------. (1969). *Speech Acts*. Cambridge, Cambridge University Press.

Sellars, W. (1963). 'Philosophy and the Scientific Image of Man'. *Science Perception and Reality*. London, Routledge and Kegan Paul. pp. 1-41.

------------. (1997). *Empiricism and the Philosophy of Mind*. Cambridge, Harvard University Press.

Shafer-Landau, R. (2003). *Moral Realism*. Oxford, Oxford University Press.

Shea, J. O. (2008).

Shoemaker, S. (2004). 'Causality and Properties' in *Metaphysics*, (eds.) J. Kim and E.Sosa. Oxford, Blackwell Publishing. pp. 253-269.

Sidgwick, H. (1967). *The Methods of Ethics*. London, Macmillan and Company.

Sinnott-Armstrong, W. (2007). 'Expressivism and Embedding'. *Foundations of Ethics*. R. Shafer-Landau and T.Cuneo. Oxford, Blackwell Publishing. pp. 485-495.

Smith A. (2007). *The Theory of Moral Sentiments*. Cambridge, Cambridge University Press.

Smith, M. (1994). *The Moral Problem*. Oxford, Oxford University Press.

-----------. (1997). 'In Defense of The Moral Problem: A Reply to Brink, Copp and Sayre-McCord'. *Ethics* 108. pp. 84-119.

Sosa, E. (1991). 'Reliabilism and Intellectual Virtue' in *Knowledge and Perspective* Cambridge, Cambridge University Press. pp. 131-145.

----------. (2001). 'Epistemology and Primitive Truth' in *The Nature of Truth*, (ed.)M.Lynch. Hong Kong , The MIT Press. pp. 641-663.

----------. (2007). *A Virtue Epistemology, Vol.I*. Oxford, Oxford University Press.

----------. (2008). 'The Raft and the Pyramid' in *Epistemology,* (eds.) E.Sosa, J.Kim, J.Fantl and M.MacGrath. Oxford, Oxford University Press. pp. 145-165.

Stanley, J. (2005). *Knowledge and Practical Interests*. Oxford, Oxford University Press.

Stevenson, C. L. (1937). 'The Emotive Meaning of Ethical Terms'. *Mind* Vol. 46(181). pp 14-31.

Stich, S. (1998). 'Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity'. *Rethinking Intuition,* (ed.) M. DePaul and W.Ramsey. London, Rowman and Littlefield Publishers. pp. 95-112.

Strawson, P. (1987). *Scepticism and Naturalism*. London, Methuen.

Sturgeon, N. (2007). 'Moral Explanations'. *Foundations of Ethics*, (eds.) R. Shafer-Landau and T.Cuneo. Oxford, Blackwell Publishing. pp.337-353.

Szabo, Z. G. (2008). 'The Distinction Between Semantics and Pragmatics'. *The Oxford Handbook of Philosophy of Language*, (eds.) E. Lepore and B. C. Smith. Oxford, Oxford University Press. pp. 361-393.

Timmons, M. (1999). *Morality Without Foundations*. Oxford, Oxford University Press.

Travis, C. (2008). 'Psychologism'. *The Oxford Handbook to Philosophy of*

Language, (eds.) E. Lepore. and B.C.Smith. Oxford, Oxford University Press. pp. 103-27.

Vahid, H. (2005). *Epistemic Justification and the Sceptical Challenge*. London, Palgrave Macmillan.

Weber, M. (1982). 'Science as a Vocation'. pp. 129-159.

------------. (1982). 'The Intellectual Sphere'. pp. 350-358.

Both in *From Max Weber*, (eds.) H. H. Gerth and C.W.Mills. London, Routledge and Kegan Paul.

Wedgwood, R. (2007). *The Nature of Normativity*. Oxford, Oxford University Press.

------------------. (2009). 'Contextualism about Justified Belief'. *Philosopher's Imprint*. Vol. 8(9). pp. 1-20.

Wiggins, D. (1999). 'Meaning and Truth-Conditions: from Frege's Grand Design to Davidson's' in *A Companion to the Philosophy of Language*. B. Hale and C.Wright. Oxford, Blackwell Publishing. pp. 3-29.

Williams, B. (1973). 'Deciding to Believe' in *Problems of the Self*. Cambridge, Cambridge University Press. pp. 136-152.

Williams, B. (1993). *Ethics and the Limits of Philosophy*. London, Fontana Press.

Williamson, T. (2000). *Knowledge and its Limits*. Oxford, Oxford University Press.

------------------. (2004). 'Philosophical 'Intuitions' and Scepticism about Judgement'. *Dialectica* 58(No.1): 109-153.

Wittgenstein, L. (2001). *Tractatus Logico-Philosophicus*. London, Routledge Classics.

--------------------. (1953). *Philosophical Investigations*. Oxford, Basil Blackwell. Translated by G.E.M. Anscombe.

Wright, C. (1991). 'Scepticism and Dreaming: Imploding the demon'. *Mind* Vol.87(1). pp. 87-116.

Zagzebski, L. (1996). *Virtues of the Mind*. Cambridge, Cambridge University Press.