

# The Primacy of Knowledge:

*A Critical Survey of Timothy Williamson's  
Views on Knowledge, Assertion and Scepticism*

**Heine A. Holmen**

Thesis submitted for the Master of Arts Degree, May 2007

Dept. of Philosophy, Classics, History of Arts and Ideas

University of Oslo



## **Acknowledgements**

Writing a thesis is a social phenomenon, so there are a number of people who deserve my innermost gratitude. As Tyler Burge once said, “no man’s intentional mental phenomena are insular”. What is thought and done in the present thesis is no exception.

First and foremost I would like to thank my supervisor Carsten Hansen for his patient support and crucial encouragement when it was most needed. Without the benefit of his insight and his capacity for sharply locating what is essential in a vast and complex field of philosophical research, very little of this essay would have seen the light of day.

Furthermore, I have been fortunate to benefit from conversations with Robert Brandom, John Hawthorne, and Timothy Williamson, when they have been visiting the University of Oslo. I am particularly grateful to Shaun Nichols for sharing his insight from the frontiers of experimental philosophy with me. I must also thank Herman Cappelen - whom got me started.

Last, but not least, I have benefited from many friends and fellow students in philosophical discussion over these and related issues, and I would like to thank them all. Above all, I would like to thank Torbjørn Gundersen, Torfinn T. Huvenes, Marianne Vahl, and Anders Strand, who have read drafts and given me invaluable input on this thesis. An especially warm gratitude goes to Trine Antonsen, for support, compassion, and a shared love of wisdom.

*“Man’s respect for knowledge is  
one of his most peculiar characteristics”*

Imre Lakatos

# CONTENTS

<b>INTRODUCTION: WILLIAMSON’S COGNITIVE TURN.....</b>	<b>7</b>
<b>CHAPTER 1: KNOWLEDGE FIRST!.....</b>	<b>8</b>
1.1 THE KNOWER AND THE CONTENT KNOWN .....	8
1.2 THE TRADITIONAL ANALYSIS.....	12
1.3 THE GETTIER PROBLEM.....	13
1.4 KNOWLEDGE FIRST! .....	15
<b>CHAPTER 2: EPISTEMOLOGY AS PHILOSOPHY OF MIND</b>	
<b>.....</b>	<b>17</b>
2.1 KNOWLEDGE AS A MENTAL STATE.....	17
2.2 THE LOSS OF SELF-KNOWLEDGE .....	22
2.3 A PROBLEM ABOUT BELIEF .....	27
2.4 TWIN EARTH .....	30
2.5 FACTORISATION AND THE INTERNALIST DIVIDE AND CONQUER.....	34
2.6 THE CAUSAL NEXUS .....	38
2.7 CONCLUSION .....	40
<b>CHAPTER 3: KNOWLEDGE AS JUSTIFICATION .....</b>	<b>41</b>
3.1 EVIDENCE.....	41
3.2 OBJECTIONS AND REPLIES TO THE EQUATION .....	47
3.3 INEXACT KNOWLEDGE AND MARGINS FOR ERROR .....	49
3.4 CONCLUSION .....	51
<b>CHAPTER 4: ASSERTION .....</b>	<b>52</b>
4.1 ILLOCUTIONARY ACTS AND CONSTITUTIVE RULES.....	53
4.2 THE STRUCTURE OF A SIMPLE ACCOUNT .....	55
4.3 THE TRUTH ACCOUNT.....	57
4.4 THE KNOWLEDGE ACCOUNT .....	61
4.5 AN ARGUMENT FROM PROBABILITY.....	62
4.6 CONVERSATIONAL PATTERNS.....	65
4.7 MOORE’S PARADOX.....	67
4.8 CONCLUSION .....	68
<b>CHAPTER 5: SCEPTICISM .....</b>	<b>69</b>
5.1 SCEPTICISM AND ITS DATA .....	69
5.2 THE STRUCTURE OF THE CLASSICAL ARGUMENT FOR SCEPTICISM .....	73
5.3 SCEPTICAL ARGUMENTS .....	76
5.4 UNIVERSAL IGNORANCE.....	78
5.5 WHERE IGNORANCE ENJOINS SILENCE AS WELL AS NIHILISM.....	81
5.6 CONCLUSION .....	82
<b>CHAPTER 6: ANTI-SCEPTICISM.....</b>	<b>83</b>

6.1 RATIONAL MANAGEMENT OF ONE’S EVIDENCE.....	83
6.2 DEFENDING CLOSURE .....	89
6.3 KNOWLEDGE-SHIFTING STRATEGIES.....	91
6.4 PROBLEMS WITH THE KNOWLEDGE-SHIFTING STRATEGIES .....	97
6.5 THE ILLUSION OF IGNORANCE.....	100
6.6 THE POSSIBILITY OF EMPIRICAL SUPPORT FOR THE ILLUSION OF IGNORANCE .....	103
6.7 KNOWLEDGE AND PRACTICAL REASONING .....	105
6.8 CONCLUSION .....	107
<b>BIBLIOGRAPHY: .....</b>	<b>109</b>

## Introduction: Williamson's cognitive turn

In *Knowledge and Its Limits*, as well as in a number of papers, Timothy Williamson develops a new, 'knowledge first' approach to epistemology and philosophy, where the notion of knowledge plays *the* leading role. His proposal is to take knowledge as a basic and primitive notion, from which we explain other important concepts and distinctions in our conceptual scheme. Philosophical orthodoxy has for a period of time taken belief to be the central explanatory notion in the conceptual triangle involving assertion, truth and action, but on Williamson's conception there is a change in roles, and knowledge is put to do explanatory work on the triangle as well as on the notion of belief itself.

The primitiveness of knowledge does not preclude reflective understanding, since we can elucidate the notion of knowledge through its role in our thinking and by providing a *modest positive account*. Williamson builds such an account on his core thesis that knowledge is a *sui generis* mental state. According to Williamson, the theory of knowledge can be firmly established as a branch in the philosophy of mind. In this sense, his 'knowledge first' epistemology takes a cognitive turn. Williamson makes this foundation his point of departure from which he can develop non-circular accounts of evidence, justification, warranted assertibility, and elucidate practical reasoning as well as refute scepticism. Those accounts employ knowledge as their chief explainer, and in return they can help elucidate the notion of knowledge by revealing its role in our thought. Hence a two-way explanatory relation unfolds itself: By placing knowledge at the centre of epistemology, and situating epistemology in the philosophy of mind, Williamson can provide rigorous accounts of central notions in our conceptual scheme. In return, those accounts provide material for a reflective understanding of knowledge that compensates for our lack of analysis. What's more, the correctness of those accounts will provide a strong justification for Williamson's conception of knowledge.

For this reason, an assessment of Williamson's approach will need to consider both whether his conception of knowledge can provide an account of these notions, and whether those explanations can help us understand knowledge itself as well as justify this conception. In light of this, the present thesis starts with a description of Williamson's account of knowledge. First, we observe the unanalysability of knowledge (chapter 1), before we encounter knowledge as a mental state (chapter 2), and as the basis for justification (chapter 3). The rest of this thesis explores two of Williamson's main applications of his account: the elucidation of assertion (chapter 4), and the refutation of scepticism (chapters 5 and 6).

# Chapter 1: Knowledge first!

*If I had to summarize this book in two words, they would be: knowledge first.*

Timothy Williamson

Famously, Edmund Gettier (1963) challenged the adequacy of the traditional analysis of knowledge as justified true belief by providing some intuitive counterexamples. A consequence of this blow was a period of intense *gettierization* in epistemology, where philosophers focused their intellectual energy on providing satisfactory conditions for knowledge. On that strategy, philosophers tried to alter the old conditions or add new ones in order to immunize the definition of knowledge from counterexamples. According to Williamson, the main motive behind this strategy is to sustain philosophical orthodoxy, where knowledge is conceived as a peculiar kind of true belief. This conception has encouraged a decomposition of ‘knowledge’ into a compound of conceptually prior notions, such as ‘belief’, ‘justification’ or ‘truth’ (2000a, p. 2). Arguably, the strategy is a failure, and so Williamson recommends that we take on a different programme in epistemology, which involves a reversal of the order of conceptual priority. The new strategy takes knowledge as a *primitive notion* that *cannot* and hence *need not* be specified in other terms. Important explanations, such as the difference between knowledge and ignorance, or the relationship between knowledge and belief, must take knowledge as their starting point in terms of which one can explain other things. In other words: *Knowledge first* (2000a, p. v).

We begin in section 1.1 by discussing two constituents of the *analysandum*, whereas we turn to the third constituent and the analysis of knowledge as a *justified true belief* in section 1.2. In section 1.3 I illustrate the *Gettier problem*, while in section 1.4 I turn to the suggestion that the *unanalysability* of knowledge should be our working hypothesis.

## 1.1 The knower and the content known

Epistemologists have long sought to analyse knowledge, motivated by a desire to understand it more fully. Their objective has been to state conditions that are individually necessary and jointly sufficient for *propositional knowledge*, which we can represent as a *canonically formed* universally quantified biconditional: For all  $x$ ,  $x$  is knowledge iff  $C(x)$  (Beaney 2003; King 1998, p. 155). The *analysandum* are instances of the schema “S knows that  $p$ ”, where ‘S’ refers to the *epistemic subject* and ‘ $p$ ’ refers to the *content known*, while  $C(x)$  is the complex expression that defines knowledge (Steup 2006). In this section I will



discuss the first two constituents of the analysandum, before we discuss the knowledge relation in the subsequent sections. Let us start with the epistemic subject (S).

It is a controversial matter to decide what kind of creatures qualify as epistemic subjects and knowledge-bearers. For instance, it is in accordance with ordinary linguistic practice to attribute knowledge to animals, since sometimes we say things like a dog *knows* when it is time for a walk (Steup 2006). However, the semantic implications are controversial.<sup>1</sup> Robert Brandom represents a line of thought that takes knowledge to be essentially a social status restricted to the kingdom of sapience, and that one cannot extend it to cover merely sentient creatures unless one applies the term by courtesy of the word (Brandom 1994, p. 5; 1995, p. 424). On his view, knowledge may be rightfully attributed to S when and only when we can adopt certain attitudes by attributing to S both: (a) a *commitment*; and (b) an *entitlement* that *p*; at the same time as the attributor (c) *undertakes* a commitment that *p* (1995, p. 429). For Brandom knowing takes place in *the space of reasons*, which one can only enter by partaking in the game of giving and asking for reasons. That move disqualifies most animals, since they cannot undertake commitments or secure entitlements (1995, p. 429).<sup>2</sup> To avoid controversy I will only speak of paradigmatic epistemic subjects such as the rational human agent. Let us now turn to the content known.

It is common to cash out the content of what is known as a *proposition* or *propositional content*, and as a result knowing is grouped among the *propositional attitudes* (PAs) (Searle 1979, p. 1; Steup 2006). Propositions are what we believe, know, and desire, but also what we assert, disconfirm, and assent to (Quine 1960, p. 192).<sup>3</sup> Accordingly, we commit to propositions as entities in our ordinary, prephilosophical speech. They are the bearers of truth and falsity, the objects of our attitudes and assertions, and they serve as premises and conclusions of inferences (Brandom 1995, p. 431n1; Soames 2002, p. 28, 131; 1999, p. 14). Propositions also seem to be causally relevant, since whether S believes that, “Ségolène Royal is the socialist president candidate”, or, “Ségolène Royal is a Hungarian trapeze artist”, makes a causal difference. For one thing, S would provide a different answer and so produce different sounds when asked, “Who is Ségolène Royal?”

This last point may prove to be a problem, because propositions are abstract objects, and, ever since Frege took propositions to be denizens of a pseudo-Platonic “third realm”,

---

<sup>1</sup> Dretske takes these attributions to be a main motive behind his epistemology, as he says: “I wanted a characterization that would at least allow for the possibility that animals (a frog, rat, or my dog) could know things without my having to suppose them capable of the more sophisticated intellectual operations involved in traditional analyses of knowledge” (1985, p. 177).

<sup>2</sup> Another point is that knowing entails believing, and one cannot believe *p* unless one *understands* it, which arguably requires a sophisticated intellectual effort that extends the capacities of merely sentient creatures (Brandom 1995, p. 425).

<sup>3</sup> Russell defines propositions as: “What we believe when we believe truly or falsely” (1919b, p. 285). On his view beliefs relate the believer to propositions that are either true or false, which means that either they refer to a fact (true) or not (false) (1919b, p. 285).

philosophers have experienced “nominalistic qualms” regarding these objects (Frege 1918; Perry 2001, p. 20).<sup>4</sup> According to John Perry, the real question is whether propositions can be intrinsically involved in thoughts as determiners of their causal relations without having causal powers themselves. Perry thinks that an affirmative answer to this question sounds suspiciously like a philosophical *deus ex machina*, so he suggests that we take a *naturalistic tack* on propositions. On that approach they are abstract objects with no causal powers of their own, but which nonetheless can be used to identify or classify representations or cognitive states that have causal powers (2001, p. 20).<sup>5</sup> If we are to be consistent with Perry’s naturalistic approach, the causal relevance of propositions must be mediated. As a consequence, the attitudinal- and assertive relations are mediated as well, and the agent is required to *do* something in virtue of which propositions become causally relevant (Soames 2002, p. 131). Let me illustrate: We can say that a *statement* is *that which one states*, whereas the particular act of stating a content is something that happens at a certain place and time. Hence it is only the latter that is part of the causal nexus (Soames 1999, p. 14). We can identify the statement is identified with “what is said” or stated, which is the proposition expressed. What one does in order to express that statement is to make an utterance, and so utterances can mediate a relation between a speaker and a proposition (Hawthorne 2004, p. 51; Soames 1999, p. 14).

We can also ask what the nature of a proposition is and query about its existence. Admittedly, that outruns the present thesis, but let me sketch something I take to be on the right track: It seems that propositions are something over and above utterances (sentence *tokens*), sentences (sentence *types*), and meanings (semantic value) (Soames 1999, p. 14). We can prove that negatively by rejecting the proposed identifications. Suppose two speakers, *x* and *y*, assert the same proposition. Then evidently they do so by producing different utterances. After all, utterances are tokens. Since the utterances are different, while the statements are the same, *x*’s utterance cannot be identical with what *x* stated, or *y*’s utterance cannot be identical with what *y* stated, or both (Soames 1999, p. 15):

- (1)  $U_x \neq U_y$
- (2)  $P_x = P_y$
- (3)  $\sim(U_x = P_x \ \& \ U_y = P_y)$

---

<sup>4</sup> Davidson calls them “intensional entities”, and he is generally hesitant to whether they can do explanatory work in a theory of meaning (1968, p. 99). His reluctance is related to his commitment to an essential indeterminacy in all interpretation (and translation). Such indeterminacy is in tension with an entification of meaning, since that would invite sharpness and go against indeterminacy (1968, p. 101).

<sup>5</sup> Perry thinks of propositions as “analogous to weights and lengths”, and the states they classify are “analogous to the complex concrete physical factors that constitute an object’s having the weight or length that it has” (2001, p. 21).

The question is whether it is plausible that  $x$ 's utterance is identical with the proposition  $x$  asserted, while  $y$ 's utterance is not identical to the proposition  $y$  asserted. If so, then by (2)  $x$ 's utterance is identical to the proposition that  $y$  asserted, while  $y$ 's own utterance is not identical with what  $y$  asserted. Consequently  $y$  asserted something that  $y$  did not produce, which seems odd (Soames 1999, p. 15). This suggests an additional premise, which takes it that if  $U$  is an utterance of a proposition  $P$ , then  $P$  is identical with an utterance only if it is identical with  $U$  (4). That validates the following entailment (4) from (1), (2) and (5):

$$(4) (u)(s) (uRs \supset (u') (s = u' \supset s = u))$$
<sup>6</sup>

$$(5) U_x \neq P_x \ \& \ U_y \neq P_y$$

Since this argument can be produced for any  $U$ s and  $P$ s the argument generalises, and we can conclude that no proposition can be identified with an utterance. Essentially, the same argument can be deployed to distinguish between propositions and sentences, since the same proposition can be asserted by uttering different sentences (Soames 1999, p. 15). I may utter, "London is pretty", and assert the same proposition, as does "puzzling" Pierre, when he says, "*Londres est jolie*".<sup>7</sup> These are not identical sentences, so we can run the argument all over and prove that no proposition is identical to a sentence.

To take down the third proposed identification an analogous argument can be used to show that propositions are not identical with *meanings* (or *semantic values*), since we may express the same proposition with different meanings. Suppose I utter, "Pierre is French", while Pierre says, "I am French". We obviously asserted the same proposition, but we did that by expressing different meanings (Soames 1999, p. 16).<sup>8</sup> Perhaps we can think of meanings or semantic values as functions from contexts of utterance to propositions (semantically) expressed (whether propositions are conceived as sets of possible worlds or structured propositions), and hence that meanings are related to propositions as functions are to their values (Hawthorne 2004, p. 51-2; Soames 1999, p. 17). If that is the true nature of the relationship between meaning and propositions, it is clear that we can always generate two different functions that relate to the same proposition, since we can always construct different functions that relate in a distinct manner to the same nonempty set of propositions (Boolos, Burges and Jeffrey 2002, p. 5). We may therefore conclude that we have a theoretical need for

---

<sup>6</sup> 'R' stands for the relation (unspecified) that holds between utterances and propositions asserted by those utterances (Soames 1999, p. 15).

<sup>7</sup> This example is taken from Kripke (1979, p. 414).

<sup>8</sup> See Soames (1999, p. 16-7). We often call the former an *eternal sentence* (i.e. its meaning expresses the same proposition in every context of utterance), and the latter an *occasion sentence* (i.e. that can express different propositions in different contexts) (Soames 1999, p. 17).

propositions (whatever their nature turns out to be) as something over and above utterances, sentences and meanings.<sup>9</sup> Let us now turn to the third constituent - the knowledge relation.

## 1.2 The traditional analysis

A traditional analysis of the schema “S knows that  $p$ ” is the JTB Analysis of Knowledge. On JTB, S knows  $p$  if and only if S has a *justified true belief* that  $p$  (Dancy 1985, p. 23; Gettier 1963, p. 58; Steup 2006). We can formulate JTB in canonical form accordingly:

(JTB)  $(\forall x) (x \text{ knows } p \equiv (p \text{ is true} \ \& \ x \text{ believes } p \ \& \ x \text{ is justified in believing } p))$ <sup>10</sup>

The left side of the biconditional is the *analysandum*, whereas the right side represents the *analysans*, which is the list of conditions individually necessary and jointly sufficient for knowledge, which are called the justification-, truth- (factivity) and belief condition (Shope 1983, p. 3). Gettier cases aside, there have been disputes regarding the individual necessity of the JTB-conditions. To demonstrate the failure of individual necessity one must provide a case where S knows  $p$  without satisfying all the conditions. I will not try to establish conclusively any of these conditions, but let me address some standard counterexamples and replies to illustrate the debate.

On the “founding insight” of *reliabilism*, there are cases where S knows  $p$  without justification (Brandom 2000, p. 97). According to the reliabilist, what make true beliefs knowledge are not the presence of justification, but the *reliability* of the cognitive processes forming the belief. As long as those processes are reliable *enough* S knows  $p$  (Steup 2006). That position is somewhat perplexing, since there is no “principled reason not to count thermometers as knowers”, and arguably that makes the position “a bit cavalier” (Brandom 1995, p. 425). Another thing is that the position offers no advantage vis-à-vis the Gettier cases, since we can make analogue cases that take down the reliabilist analysis of knowledge, but we will return to this below.

---

<sup>9</sup> An instance of the schema “S knows that  $p$ ” is used to *ascribe* knowledge to S, a rational human subject, by uttering  $P$  in a context C. The utterance employs  $P$ ’s semantic value to semantically express a proposition, which, if the ascription succeeds, will attribute knowledge of some proposition  $p$  to S. The attribution of knowledge is conceived as implicitly relating a time  $t$  and a possible world  $w$  to the subject of the ascription, so the attribution to S expresses the proposition  $K[S,p,t,w]$  (Williamson 2005a, p. 214). If S knows  $p$  at  $t$  in  $w$ , then it is true that  $K[S,p]$  at the given time  $t$  in the given possible world  $w$ . Think of a *possible world* as just a way the universe could have been, so if  $p$  is said to be true in  $w$   $p$  would have been true if  $w$  obtained (Soames 2002, p. 23). This talk about possible worlds is the non-committed and metaphysically speaking innocuous and non-reductive approach to possible worlds and modal discourse, which was advocated by Saul Kripke (1981, p. 44-5). On this view talk, about possible worlds may serve to illuminate modal discourse without providing a reductive analysis of it (Soames 2002, p. 23). As a consequence, possible worlds are considered to be things that are stipulated in a sense that preserves trans-world identity of at least some of its inhabitants. An attempt at reductive analysis is the renowned *modal realism* approach that was advocated by David Lewis (1968). On that conception possible worlds were discovered rather than stipulated, and the inhabitants were counterparts rather than identical to each other.

<sup>10</sup> Alternative formulations are: (1) ( $p$  is true & S believes  $p$  &  $p$  is *evident* to S) (Klein 1971, p. 61); (2) (S has *adequate evidence* for  $p$ ) (Chisholm); (3) (S is *sure* that  $p$  is true) (Ayer). Gettier takes them to say the same thing, i.e. that S must be *justified* (1963, p. 58).

The belief condition has also been contested with putative counterexamples where S knows  $p$  without believing it. Colin Radford has argued for this with a case of the unconfident student, Jean, who (to his own surprise) provides many correct answers during an examination in English history (1966). Since Jean sincerely takes himself to be ignorant about English history he does not believe that his answers are correct. Yet, he seems to know a lot of English history (Radford 1966; Radford 1988, p. 496; Steup 2006; Williamson 2000a, p. 42). It is not entirely clear whether Jean knows beforehand or comes to know during the process of examination, and so the answers may just be the results of “pure fluke”? (Radford 1988, p. 498) However, it seems extreme to assert flat-out that Jean does not know any English history after a performance like that, so it seems more reasonable to dispute that Jean does not believe. Arguably, Jean is in a state belief-like enough to be counted as believing when he provides answers to questions like: “When was the battle of Agincourt?”

The truth-condition of JTB enjoys almost universal consent among epistemologists (Steup 2006),<sup>11</sup> even though some philosophers of science have disputed the condition (Cartwright 1980; Longino 2002). To retract the truth condition on knowledge is extremely awkward, and it would be very odd if someone could say something like, “I know  $p$ , yet  $p$  is false”. Since most epistemologists accept that condition, I will largely ignore the discussion that is raised by these philosophers. We will return to these issues and perhaps strengthen their cause when I flesh out Williamson’s position below.

### 1.3 The Gettier problem

The *Gettier Problem* is a different challenge to JTB, and seeks to undermine the claim that the three aforementioned conditions are *jointly sufficient* for knowledge. It does not question their individual necessity. The following section will illustrate the debate about this problem. The rationale behind Gettier’s counterexamples is to describe a situation where JTB conditions are satisfied, but where S does not know  $p$ . The possibility of such a scenario will demonstrate that JTB conditions are not jointly sufficient for knowledge (Gettier 1963, p. 58).

Suppose that you are justified in believing *falsely*<sup>12</sup> that Bjørn’s legs are paralysed (6), since he behaves as if he suffered from a severe handicap.<sup>13</sup> Perhaps you have seen Bjørn in a wheelchair behaving in a manner indistinguishable to the untrained eye from other disabled people, or he has shown you his welfare card and facilities back home. Maybe the doctor is in

---

<sup>11</sup> Aristotle endorses this view when he says, “[W]e cannot know what is not true” (2005b, p. 679).

<sup>12</sup> Gettier: “[I]t is possible for a person to be justified in believing a proposition which is in fact false” (1963, p. 58).

<sup>13</sup> The example (analogous to Gettier’s “Case II” (1963, p. 59)) is from a novel by Norwegian writer Dag Solstad, where the main character, Bjørn Hansen, fools his family and friends into believing that he suffers from a severe handicap (1992, p. 140-4, 132-8).

on it too, and has assured you of Bjørn's condition. All this can warrant a justified belief that (6) is true. But from (6) you are also justified in believing (7):

(4) Bjørn's both legs are paralysed.

(5) Either Bjørn's both legs are paralysed or Pierre (another friend) is in San Tropez.

Unbeknownst to you, Pierre is in fact in San Tropez. Since a true disjunction only requires one true disjunct, (7) is in fact true. Hence (7) is knowledge on the JTB Analysis since it is a *true* justified belief. But intuitively (7) is not something you know. The fact that you have formed a *true* belief (7) seems to be the result of *pure luck* and independent of your justification. Let us say that, *if* beliefs are true independently of their justification, then they result from *epistemic luck* (Engel 2002, p. 3-4; Steup 2006). Then our intuition is that knowledge is incompatible with epistemic luck. As a result, (7) provides a counterexample to JTB since it does not have the "appropriate justification-relation" (Engel 2002, p. 4).

Note that the basis on which we justify (7) is that justification is closed over justified entailments. Thus the proponents of these Gettier-cases endorse something like:

(Closure (JB))                      If one has justified belief in  $p_1, \dots, p_n$  and believes  $q$  in the light of competent deduction  $q$  from  $p_1, \dots, p_n$  then one has a justified belief that  $q$ .<sup>14</sup>

If one accepts *Closure (JB)* (CJB) one can competently infer and justify (7) from the justified belief (6). Importantly, *CJB* does not require the *truth* of  $p_1, \dots, p_n$  in order for them to justify  $q$ . In other words, one can justify (7) through a justified *false* belief (6).

We can also make an analogous case against the reliabilist analysis of knowledge (RAK). Recall that on RAK a true belief is knowledge only if a reliable cognitive process forms it. Sense perception is normally taken to be such a reliable process, so to see  $p$  when  $p$  is true will constitute knowledge on RAK (Steup 2006). Now a version of Goldman's famous *Barn County* case will suffice to make a counterexample to reliabilism (1976): Suppose you drive through a county peppered with barns, and you look at them from your car. You may reliably form the belief that, "There is a barn nearby". Since that is true and your sense perception is reliable, that belief will constitute knowledge on RAK. But suppose that you drive past the one and only *fake barn* in Barn County, which has a *fake façade* (i.e. a structure that from the road look exactly like a real barn). Then your sense perception will be

---

<sup>14</sup> See Williamson (2005b, p. 687). Hawthorne's critique of *Intuitive closure* (IC) can also be applied to *CJB*. Suppose that you at  $t_1$  is justified that  $p_1, \dots, p_n$  and you commence a lengthy deduction that establishes  $q$  at  $t_3$ . Suppose further that at  $t_2$  you loose justification of  $p_1, \dots, p_n$  for some reason (perhaps due to misleading evidence). Then intuitively the connection is broken so that  $q$  cannot be justified at  $t_3$ . In order to avoid this we must require the agent to retain her justification *throughout* the deduction. See Hawthorne (2004, p. 32-3).

responsible for you forming the false belief that, “There is a barn nearby”. From that belief you may infer that, “Either there is a barn nearby, or Pierre is in San Tropez”. But Pierre *is* in San Tropez, so this is a true belief formed by a reliable cognitive process, and hence counts as knowledge on RAK. The problem is that intuitively you do not know this, since it is only by epistemic luck that you hold that true belief. Hence this is a counterexample to reliabilism.

Let this be enough to illustrate the debate about the Gettier problem. The debate has become an exercise where philosophers provide ever-sophisticated accounts only to be met with ever-sophisticated counterexamples. Shope says that the discussion in his book can serve as a “sustained illustration” of “the role that is often played in contemporary philosophical debate by the technique of seeking counterexamples” (1983, p. xi). I think he’s right, and this fact may indicate that something is wrong about the choice of strategy.

#### **1.4 Knowledge first!**

The fact that many eminent philosophers over the past thirty years have invested so much work and intellectual effort on an analysis of knowledge without succeeding indicates that something is wrong with their strategy. I will end this chapter by considering some reasons that *prima facie* favour analysis, but which nevertheless does not necessitate this strategy. Then I turn to Williamson’s alternative.

It is often held that we can expect an analysis of knowledge in terms of belief since knowledge entails belief but not vice versa (Williamson 2000a, p. 3). But such entailments do not necessitate analysis in terms of what is entailed, since, even though red entails being coloured, we do not expect a non-circular analysis of red that conjoins colour with other properties specified without reference to red (2000a, p. 3, 32). We can perhaps approximate an analysis of knowledge based on belief, but that too does *not guarantee* analysis (2000a, p. 4, 33). For instance, a good approximation of parenthood is that x is a parent of y if and only if x is an ancestor of y and x is not an ancestor of an ancestor of y. By considering *recherché* cases of incest we can see that this approximation can never take us to a complete analysis. Consider King Oedipus’ tragic fate:<sup>15</sup> Oedipus’ mother, Jocasta, begets no less than four children with her son before they discover the truth. For these children there is a parent (Jocasta) that is both a parent to the children as well as an ancestor to one of their ancestors (Oedipus), which is a crucial failure on the definition.

Admittedly, none of these replies are *conclusive*, since they only show that knowledge is *not exceptional* if it is not analysable through its entailments (Williamson 2000a, p. 32-3).

---

<sup>15</sup> See Sophocles’ classical Greek playwright *Oedipus Rex* (Sophocles 1992, p. 176-221).

We can conclude that an analysis is *not* compulsory, though it remains a possibility. Importantly, that option has become *less plausible* over the years when we consider the vast inductive basis from a long history of failed analysis (Shope 1983). Inductions are also inconclusive evidence,<sup>16</sup> but they can strengthen the suspicion that analysis of knowledge is a “degenerating research programme” in contemporary philosophy (Williamson 2000a, p. 31). Williamson’s basic conviction is that the failure of more or less sophisticated definitions *need not* encourage us to reanalyse knowledge, and that Gettier’s results can be seen from a different perspective: Perhaps they are signs of the conceptual primacy of knowledge?

As a response, Williamson denies the need for and the plausibility of a complete analysis of knowledge. Let us label this the *Unanalysability Thesis* (UT). Since the prospect of analysis is slim, Williamson’s proposal is to take UT as our *working hypothesis* and not anticipate an analysis of knowledge (2000a, p. 33). We may still get a reflective understanding of knowledge by providing a “thinner” modest positive account (MPA) of the concept, which elucidates what relations knowledge has to other concepts in our conceptual scheme. Hence we must try to get a grip on the central role of knowledge in our thought (DeRose 2002, p. 573; Williamson 2000a, p. 33-4). Accordingly, we may expect knowledge to be elucidated when we see its role vis-à-vis other important notions, such as truth, belief, justification, assertion and ignorance. The rest of this thesis seeks to understand the role of knowledge as an explainer of these important notions, and to see what new light these explanations may throw back on knowledge itself, and also to what extent the utility of those accounts may serve to justify Williamson’s conception of knowledge.

---

<sup>16</sup> After all it is an induction so if a complete analysis of knowledge in terms of belief should turn up tomorrow no contradiction would follow between that event and the inductive argument. This is what David Hume discovered in his problem of induction (1993, p. 22-3).



## Chapter 2: Epistemology as Philosophy of Mind

*We can see epistemology as a branch of the philosophy of mind.*

Timothy Williamson

The pivotal claim in *Knowledge and Its Limits* is that knowledge is a mental state of a mind that is deeply entrenched in the world external to the agent's physical boundaries. As a consequence of the first claim, Williamson takes epistemology to be a branch in the philosophy of mind, while the second commits him to a radical form of externalism. It seems that the only one way for knowledge to be a purely mental state is if what goes on in the mind constitutively depends on environmental factors. The idea is to consolidate our understanding of knowledge through the philosophy of mind. We start in section 2.1 with a characterisation of the position, where I present Williamson's *modest positive account* of 'knowledge'. The rest of this chapter is a critical scrutiny of his position. In section 2.2 I discuss whether transparency (or luminosity) is a mark of the mental, while in 2.3 I explore the relationship between knowledge and belief. Section 2.4 offers initial defence for general externalism, while sections 2.5 and 2.6 discuss objections to the idea of knowledge as a state of mind.

### 2.1 Knowledge as a mental state

In order to provide a foundation for his account of knowledge, Williamson commits to the existence of a mental state that is necessary and sufficient for knowing; our *initial question* should therefore be whether such a state exists (Williamson 2000a, p. 21). It is not obvious that such a state should *not* exist. After all, knowing is a *propositional attitude* (PA), and PA-ascriptions are a primary means to identify mental states; so that when one attributes knowledge to some agent (S) one obviously attribute some mental state to S (Burge 1979, p. 599). However, Williamson's commitment is more radical, since he commits to a *purely* mental state *M* that is necessary and sufficient for knowing. That *is* radical because knowledge is one of those PAs that are *factive* and entail the truth of what they relate to. When S has a *factive (propositional) attitude* (FA) to *p*, it entails that *p* is true. Hence FAs are states that constitutively depend on environmental factors.<sup>17</sup> Let us for the present purpose roughly characterise such a state as *externalist*. Many theorists take these facts about FAs and knowledge to indicate that they are not *purely* mental states, since they think of purely mental states as completely constituted by what goes on within the skull (or skin) of the individual

---

<sup>17</sup> FAs are a subset of PAs that relate only to *true* propositions (= facts). Examples of FAs are: "seeing that *p*", "remembering that *p*", and "perceiving that *p*". Importantly neither "Believing that *p*" nor "Being justified that *p*" are FAs. Aristotle seems to acknowledge the existence of such states when he says that: "Among our *intellectual states* that grasp the truth, some – knowledge and understanding – are always true, whereas others – for example, belief and reasoning – admit of being false" (2005b, p. 692).

(Williamson 2000a, p. 49). Let us roughly characterize such states as *internalist*. As a consequence of this intuition, one takes S to know *p* if and only if S is in a purely mental state *M* and some additional factor obtains that depends on the environment. That goes against Williamson's commitment to the existence of a mental state *M* that is necessary and sufficient for knowing. In this sense, Williamson's commitment is linked to our roughly articulated externalism. Let us start by getting a clearer picture of just what his position amounts to, before we return to the issue of the externalist commitments below.

Pretheoretically, the paradigmatically mental PAs resemble FAs to such a degree that it would make intuitive sense if the latter were mental<sup>18</sup> too. So unless theoretical reasons go against it, knowledge *could* count as mental (Williamson 2000a, p. 22). When one takes knowledge to be a mental state, one assimilates knowledge with paradigmatically mental PAs, like *believing* and *desiring*, and contrasts it against non-examples, like *believing truly* (2000a, p. 27). Intuitively, there is no *purely* mental state *M* necessary and sufficient for *believing truly*, while there are such states for *believing* and *desiring*. Again, the basis for this intuitive delineation seems to be the fact that *believing truly* constitutively depends on external factors, i.e. the truth of what one believes, while to *believe p* does not depend on any external factors. The question is which class knowledge resembles most.

Let us begin our query by looking more closely at the notion of a 'mental state'. Williamson wisely avoids a formal definition, but says that we can get an intuitive feel by roughly defining it "in terms of the concept *mental concept of a state*" (2000a, p. 28):

(Mental state (MS))      A state is mental if and only if there could be a mental concept of that state.<sup>19</sup>

If knowledge is a mental state, then according to *MS* there could be a mental concept of it. Conversely, if there is such a concept, then it "follows immediately" that knowing is a mental state (2000a, p. 29). Williamson's strategy is to illuminate mentality by way of a theory of the mentality of concepts (Sosa forthcoming). In light of this, our initial question would be answered affirmatively by proving that there could be a mental concept to the state of knowing. According to Williamson, we can illuminate the notion of a mental concept by saying that if a concept *C* is the conjunction of concepts  $C_1, \dots, C_n$ , then *C* is mental if and only if each  $C_i$  is mental (2000a, p. 29). Again, this is no formal definition, but Williamson takes it to be sufficient for getting a hold of the notion. But is that true?

---

<sup>18</sup> I will omit the addition of 'purely' even though in this context 'mental' is to be read in that strict sense of the word.

<sup>19</sup> Note that one can have non-mental concepts to mental states. Two concepts may share extension without both being mental (i.e. without both having the intension of a mental concept) (Williamson 2000a, p. 28).

Let us see what it says: We can begin by noting that it does *not* say that states are not mental when their content-clause has a non-mental concept. If it did, then ‘believing that grass is green’ would be non-mental since ‘grass’ is non-mental (2000a, p. 29). What it *does* say is that concepts, such as ‘believing truly’ or ‘hallucinates and such that grass is green’ are non-mental, since they have non-mental concepts (‘truly’ and ‘grass’) in their conjunction of concepts (2000a, p. 29). For the same reason, ‘having a justified true belief’ is non-mental, since ‘true’ is a non-mental conjunct in that concept. Observe that if the standard analysis of knowledge as a justified, true belief were correct, then knowledge would be non-mental since it would have a non-mental conjunct (‘true’) (2000a, p. 30). This point actually applies to most standard analyses, and hence they would be doomed *if* we could show independently that ‘know’ is a mental concept. Of course, such an analysis might be extensionally identical with knowledge, but one could not identify its intension with ‘knowledge’ since the latter is a mental concept while its analysis would not be (2000a, p. 28, 30).<sup>20</sup>

Curiously enough, the above characterization seems to render concepts like, ‘believing that  $p$  and  $a = a$ ’, to be non-mental and thus non-identical with the intension of ‘believing that  $p$ ’.<sup>21</sup> This follows since one of its conjuncts, ‘ $a = a$ ’, is not intuitively mental. Yet, these concepts are extensionally identical in *all* possible worlds, and it is hard to see by what difference the one is mental while the other is not. Another worry about this characterization is that it looks like mere stipulation to say that some concept  $C$  is mental if and only if its constituents are. It seems that we get something out of that characterization only if we have a prior grasp of the difference between mental and non-mental concepts. Perhaps the distinction is intuitive and in no need of further clarification? Yet, there is a sense in which I would have welcomed a more robust characterization that could remove the appearance of stipulation.

If *MS* is accepted, it is tempting to make a conceptual turn on our initial question, and ask whether there can be a purely mental concept of knowledge. Williamson’s proposal is that knowing is a *state* of mind, so we are exclusively interested in those FAs that are *stative* (2000a, p. 34).<sup>22</sup> By making that conceptual turn, we encounter Williamson’s modest positive account (MPA) of ‘knowledge’. The MPA is a substitute for analysis: Arguably, knowledge does not factorise as standard analysis requires, but that does not block reflective

---

<sup>20</sup> This claim hinges on a distinction between mental and non-mental vocabulary on the level of intension. This idea is analogous to what Davidson defended when he advocated the impossibility of psychophysical bridge laws between mental and physical events (1970). According to him, such a connection amounts to “changing the subject” or “deciding not to accept the criterion of the mental in terms of the vocabulary of the propositional attitudes”, and he took the existence of psychophysical bridge laws to threaten the very existence of a mental vocabulary (1970, p. 216). For Williamson, it is not enough to establish an identical extension of two concepts  $\varphi$  and  $\psi$  in order to provide a conceptual definition (i.e. identifying the intensions of  $\varphi$  and  $\psi$ ). Note that if that is correct and knowledge is mental, then a possible characterization of knowledge may avoid Gettier counterexamples and still not define ‘knowledge’ (Williamson 2000a, p. 31).

<sup>21</sup> I owe this point to Anders Strand.

<sup>22</sup> Since that distinction matters little for our present purpose I will keep on talking about FAs as if they were all stative.

understanding of the concept (2000a, p. 33). Williamson's MPA is based on the notion of a *factive mental state operator* (FMSO), since FAs are characteristically expressed in natural language as FMSOs. An FMSO ( $\Phi$ ) functions syntactically as a verb and is semantically unanalysable; hence it is *not synonymous* with a complex expression, whose meaning is composed of the meanings of its parts. To count as an FMSO  $\Phi$  must also satisfy the following (2000a, p. 34-5):

- a) Typically  $\Phi$  takes something animate as subject term and a content-clause as object.
- b)  $\Phi$  is factive (i.e. an inference from "S  $\Phi$ s that A" to "A" is deductively valid).
- c) "S  $\Phi$ s that A" attributes a propositional attitude to S.

The question is whether 'know' is an FMSO. It clearly satisfies (a), and the following consideration also shows that (b) is satisfied: We could say that, "Pierre guessed *incorrectly* that penguins could fly", but to say that, "Pierre *knew incorrectly* that penguins could fly", is clearly infelicitous.<sup>23</sup> Hence "A" does not seem to be cancelable from an utterance "S knows that A". This proves that factivity is not a conversational implicature of 'knowing', but rather something like a strict implication (Williamson 2000a, p. 35).<sup>24</sup> Then 'know' satisfies (b), since "A" can be inferred in every case where "S knows A". Another question is whether 'know' is stative: In English, stative FMSOs are marked linguistically by the impropriety of progressive tenses (2000a, p. 35). We could say something like, "Pierre is proving the incompleteness theorem", but it would be ungrammatical to say, "Pierre is knowing that Theatetus cannot fly". The linguistic impropriety of the latter indicates that 'know' is stative.

That leaves us with condition (c): Can we support that knowledge-ascriptions attribute propositional attitudes? One mark of propositional attitudes is that it follows from, "S has a PA to A", that, "S grasps the proposition that A" (Williamson 2000a, p. 25). We can show that 'know' meets this requirement, since it follows from, "Pierre knows that Olga is playing chess", that, "Pierre grasps the proposition that Olga is playing chess" (2000a, p. 36). For if, contrary to our assumption, there is a part of the proposition that is not grasped by Pierre, he cannot be said to know that proposition. Suppose Pierre uttered, "I know that Olga is playing chess, but I don't know what 'chess' is." Then it seems accurate to say that Pierre attributes the following to himself, "Pierre knows that Olga is playing 'chess' (whatever that is)." So I take 'know' to be in accordance with other propositional attitudes and satisfy (c).

---

<sup>23</sup> I use "felicitous" and "infelicitous" in the sense that Austin used these terms (1979, p. 132).

<sup>24</sup> Cancelability is a criterion of implicatures (Grice 1989, p. 44). According to Grice, one can either cancel implicatures from an utterance by using contextually marked features (i.e. the form of the utterance together with the context rules out the implicature), or one can explicitly cancel it by the addition of a clause that states or implies that the speaker has "opted out" (1989, p. 39, 44).

We have already touched upon the question of whether ‘know’ is semantically unanalysable by defending  $UT^{25}$  as a working hypothesis. As we saw, the evidence for that claim was inductive and thus inconclusive. Perhaps a better strategy would be to fight the sources of counterintuitions to  $UT$ ? A main motivation in contemporary philosophy for thinking that knowledge can be compositionally analysed, is the internalist conception of mind, so it might be a good idea to defend  $UT$  by attacking internalism. We will encounter internalism below, so until then I will take the inductive evidence as support for the semantical unanalysability of ‘know’. Then we can conclude that ‘know’ satisfies the conditions for being an FMSO. Williamson takes knowing to be the most general of the stative FAs, in the sense that if “S has a stative FA to  $p$ ” it follows that “S knows  $p$ ”. Since FMSOs are the expression of stative FAs in natural language, the parallel claim on the conceptual level is that for all  $\Phi$ s, if  $\Phi$  is an FMSO, then “S  $\Phi$ s that A” entails “S knows that A” (Unger 2002, p. 158-183; Williamson 2000a, p. 37).<sup>26</sup> The following summarize our discussion:

- (Pi) If  $\Phi$  is an FMSO, then from “S  $\Phi$ s that A” one may infer “A”.
- (Pii)\* ‘Know’ is an FMSO.
- (Piii)\* If  $\Phi$  is an FMSO, then from “S  $\Phi$ s that A” one may infer “S knows that A”.

We can now articulate a modest positive account (MPA) of ‘knowledge’: The principles marked \* characterize ‘knowledge’ uniquely “up to logical equivalence”, in the sense that one grasps the concept only by being disposed to reason according to them (2000a, p. 39-40).<sup>27</sup> The result of our discussion is therefore twofold: We have a characterization of knowledge as mental at the metaphysical level through the  $MS$  and the above characterisation of a mental concept. But we also have an MPA of ‘knowing’ at the conceptual level through the two principles  $P_{ii}$  and  $P_{iii}$  and the notion of an FMSO. In essence, Williamson’s strategy has been to provide an account of the mental by way of a theory of the mentality of concepts (Sosa forthcoming). We have seen that this characterisation appears to be a matter of pure stipulation at the metaphysical level due to the non-robust portrayal of mental concepts. Importantly, that point seems to go counter to an account of mentality “by way of a *theory* of the mentality of concepts”, since some kind of robustness seems to be a fair requirement on any theory. We also argued that the concept ‘know’ is an FMSO and that its metaphysical

<sup>25</sup>  $UT$  is the working hypothesis that knowledge is unanalysable. See chapter 1.

<sup>26</sup> Let me illustrate: *To see* is an FMSO, and so “S sees that A” entails that “S knows A” (Williamson 2000a, p. 38).

<sup>27</sup> Note that the MPA does not yield decomposition. If it did, ‘know’ would be decomposable and fail to pass as an FMSO. Note that (c) blocks ‘believes truly’ from being an FMSO, which is good, since every FMSO implies ‘knowledge’. If (c) did not block that, ‘believes truly’ would entail ‘knowledge’ and that is obviously false. The same goes for ‘justified true beliefs’. See Sosa (forthcoming).

counterpart knowing (as a stative FA) has a mental concept to it. According to Williamson, it “follows immediately” on *MS* that knowing *is* a mental state. Admittedly, theoretical obstacles can block this entailment, so Williamson does not think that he has firmly established his claims by arguing that ‘know’ is an FMSO. Rather, the MPA characterise the pretheoretic intuitive thought that knowledge is a purely mental state. Hence a natural strategy is to defend those intuitions by taking down the theoretical obstacles that go against them.

## 2.2 The loss of self-knowledge

There are some reasons that go against taking knowledge to be purely mental: For one thing, it runs counter to the traditional conception of a *privileged epistemic access* to one’s own mind. The agent’s access is often assumed to be direct and transparent, but if knowledge is a purely mental state that conception is strained since we are not always in a position to know whether we know. A second reason for doubt is the assumption of non-factorisation: It is natural to think of belief as a *factor* of knowledge since the former is a necessary condition for the latter. Finally, one can be troubled about the *nature* of a mental state like knowing: Is it not odd to talk about mental states that constitutively depend on environmental factors? Can such states fit into the *causal nexus* of psychological explanation? The rest of this chapter will deal with these issues, but let us begin with the question of self-knowledge.

One reason for not regarding knowledge as purely mental is the traditional Cartesian assumption of self-knowledge, which conceives the mental as something to which humans have privileged access (Kim 1998 p. 206; Reed 2006, p. 103). If knowledge is mental we lose some self-knowledge since one is not always in a position to decide through introspection whether one knows.<sup>28</sup> Suppose *S* believes that he knows who is prime minister of Nepal (*P*), but that unbeknownst to *S* *P* has turned false due to a governmental crisis. *S* no longer knows *P*, and yet is in no position to know that she is ignorant of *P* just by being the proprietor of her mental states. Note that one must endorse the following to get this result:

(Transparency (T\*)) For every mental state *M*, whenever one is suitably alert and conceptually sophisticated, one is in a position to know whether one is in *M*.<sup>29</sup>

If *T\** is true, knowledge and other *wide* content-states cannot be purely mental, since to know whether one has them often requires external input. Unless one wants to embrace (bad)

---

<sup>28</sup> This is a version of the problem of wide content and loss of self-knowledge: There is an apparent conflict between first-person authority and direct access versus wide content, since the latter states depend constitutively on the external environment (Kim 1998, p. 206).

<sup>29</sup> See Williamson (2000a, p. 24).

idealism one cannot take the mental to provide such input. It is tempting to assert that  $T^*$  fails since there are paradigmatic mental states that flout it. Consider the state *hoping that such-and-such is the case*: Suppose Pierre thinks that he hopes socialist candidate, Ségolène Royal, loose the French presidential election. In fact she didn't win, but yet it is *not impossible* for Pierre to feel disappointed about the result. This indicates that unbeknownst to him Pierre was really hoping for her triumph (Williamson 2000a, p. 24). Yet, *hoping* is a purely mental state and so it cannot be necessarily true that mental states are transparent.

Perhaps it would count in favour of epistemic privilege that knowing seems to require reasons while that is not the case for beliefs? (2000a, p. 25) Another thing is that knowledge seems defeasible by new information in a different manner than beliefs (2000a, p. 25). Can this favour  $T^*$ ? In my view both points are flawed. There are many beliefs that are challenged by new information. For instance, would the information that your drink contains a serious drug, which prevents you from clear thinking, seriously challenge your belief that you are entertaining a dinner party by being charming and witty. We can also put the first point aside, since it is plausible that to *rationally* believe  $p$  requires reasons just as knowledge does. In that sense, the good beliefs require reasons. A related consideration in favour of  $T^*$  is that we need the principle to explain epistemic asymmetry between one's own mental states and those of other people. But that point is not very convincing, since we may explain this asymmetry by saying that access to one's own mental states *can* be non-observational, when that can *never* be the case with other people's states. The same goes for knowledge (2000a, p. 25).

Arguably, a denial of  $T^*$  generalises since we can prove that no non-trivial mental states satisfy it (Williamson 2000a, p. 24). In general, mental states are *not luminous* and we do not always know whether we are in a given state, and so we are in a sense *cognitively homeless* (Harman 2002, p. 417; Williamson 2000a, p. 94). Let us say that a *condition* is individuated by the cases in which it obtains so that two conditions  $C$  and  $C^*$  are *identical* if they obtain in exactly the same cases. Suppose that  $S$  is in a position to know  $p$  only if no obstacle blocks  $S$ ' path to know  $p$  (2000a, p. 94-5). Now define a luminous condition as:

(Luminosity) For every case  $\alpha$ , if  $C$  obtains in  $\alpha$ , one is in  $\alpha$  in a position to know that  $C$  obtains.<sup>30</sup>

Luminosity is an expression of the traditional Cartesian conviction that the mental realm is especially accessible to human agents - our *cognitive home* (Reed 2006, p. 303). Now let us see if we can challenge that picture: Suppose  $t_0, t_1, \dots, t_n$  is a series of times at one-millisecond

---

<sup>30</sup> See Williamson (2000a, p. 95).

intervals from dawn to noon, and that  $\alpha_i$  is the case at  $t_i$  ( $0 \leq i \leq n$ ). Suppose that one at dawn feels cold, but that during the time-span till noon the condition changes and one feels warm. Suppose at  $t_i$ , between  $t_0$  and  $t_n$ , one knows that one feels cold and is sufficiently confident of that in a reliable manner necessary for knowing. Then one is expected to be almost equally confident of this feeling at  $t_{i+1}$ , since if one does *not* feel cold at  $t_{i+1}$ , one's confidence at  $t_i$  would be unreliable. We have already stipulated its reliability, so at  $t_{i+1}$  one feels cold. We can summarize this as follows (Williamson 2000a, p. 97):

- (1i) If in  $\alpha_i$  one knows that one feels cold, then in  $\alpha_{i+1}$  one feels cold.

Assume that feeling cold is luminous with the aim of performing a *reductio ad absurdum*. Consequently, the following conditional is true (Williamson 2000a, p. 98):

- (2i) If in  $\alpha_i$  one feels cold, then in  $\alpha_i$  one knows that one feels cold.

One feels cold at  $\alpha_i$  (3<sub>i</sub>), so from (3<sub>i</sub>) and *modus ponens* on (2<sub>i</sub>) we get (4<sub>i</sub>); which by a further *modus ponens* on (1<sub>i</sub>) yields (5<sub>i</sub>) (Williamson 2000a, p. 98):

- (3i) In  $\alpha_i$  one feels cold.  
 (4i) In  $\alpha_i$  one knows that one feels cold.  
 (5i) In  $\alpha_{i+1}$  one feels cold.

By repeated application on (5<sub>i</sub>) we reach the conclusion that one feels cold at noon, which is contradictory to our set-up and thus yields a *reductio*. That accomplished we can negate our assumption, and hence assert that *feeling cold* is *not* a luminous condition. Since *feeling cold* is a paradigmatic mental state our argument shows that: (a) paradigmatic mental states fail to be luminous; and (b) anti-luminosity is not enough to disqualify knowledge as mental (Williamson 2000a, p. 98). According to Keith DeRose, an attractive option to “the friends of luminosity” is to admit that feeling cold is not luminous, but make a counterattack and take feeling cold to be *weakly luminous* by endorsing the following (2002, p. 576):

(Weak luminosity (WL)) For every case  $\alpha$ , if in  $\alpha$  C *safely*<sup>31</sup> obtains one is in  $\alpha$  in a position to know that it does.

---

<sup>31</sup> We can interpret “safely” as saying that condition C is “not even close to not obtaining” (DeRose 2002, p. 576).



DeRose thinks *WL* is more apt as a criterion of mentality. As a consequence, knowledge is disqualified as mental since to know is *not* weakly luminous (2002, p. 578). Against *WL* Williamson's argument is impotent since that argument depends crucially on what happens near the borderlines of the relevant conditions (DeRose 2002, p. 576). But there is reason for suspicion: First, notice that there can be two readings of DeRose's demarcation. On one reading, *WL* demarcates at the metaphysical level of obtainings and existence. If that reading is correct, *WL* has a considerable price since its qualification "safely" will qualify existence in a way that implies dubious metaphysical commitments. Gradual existence is usually shunned in metaphysical discourse, so if that reading is correct *WL* has needless costs.<sup>32</sup> We could also read the qualification epistemologically, but then notice the following: It is very tempting to say that *C* safely obtains in a case  $\alpha$  when and only when there are no sufficiently similar cases  $\beta_1, \dots, \beta_n$  where *C* fails to obtain. That is how Williamson characterises *epistemic safety*, and he says that this is a necessary reliability-condition that needs to be satisfied by any belief if that belief is to be known (2000a, p. 147). But then weak luminosity will have the same extension as the extension of reliable (i.e. knowable) beliefs, and for that reason will a plausible reading of DeRose's demarcation turn *WL* into a rather trivial principle, which says that one is always in a position to know that *C* obtains if one's belief that *C* obtains could have been known.

The significant point in DeRose's *demarcation* is that the difference between when *C* safely obtains and not creates a border against where *C* non-safely obtains, but where, importantly, *C* still *obtains*. By demarcating weak luminosity in this manner, one can block Williamson's argument, since knowledge of what obtains on DeRose's border does not risk tension with the reliability-condition on knowledge (Reed 2006, p. 304):

(Reliable (R\*)) If in  $\alpha_i$  one is in a position to know *p*, then in every case sufficiently similar to  $\alpha_i$  one does not falsely believe that *p*.<sup>33</sup>

If luminosity is demarcated at the border between where *C* obtains and no longer obtains, then one could easily be situated in a case  $\alpha$  where *C* obtains that is sufficiently similar to a case  $\alpha^*$  where *C* does not obtain. To say, as the friends of luminosity do, that one knows whether *C* obtains in those circumstances is in tension with the reliability-constraint on knowledge. Williamson's dialectics took advantage of that in his *reductio*, since Williamson was able to

---

<sup>32</sup> After all, this is what old father Parmenides taught us, when he strictly forbade any crisscrossing between being and non-being: "For in no way may this prevail, that things that are not, are. (...) Thus it must either *fully* be or not" (Parmenides 2005, p. 38).

<sup>33</sup> This formulation I have from Reed (2006, p. 304). See Williamson (2000a, p. 100).

push the obtaining of condition *feeling cold* all the way to absurdity on the assumption of  $R^*$  and knowledge that the relevant condition *feeling warm* obtained at some  $t$ . If S knew that condition C obtained at  $t_i$ , S could come to know by  $R^*$  that C also obtained at  $t_{i+1}$ . Since one comes to know that C obtains at  $t_{i+1}$  by  $R^*$ , we can just repeat the procedure to push knowledge of C further on. In light of this, weak luminosity is really just a Cartesian regrouping. Baron Reed formulates the new luminosity as follows:

(Luminosity\*) There is at least one maximally specific way in which a condition C may obtain such that, for every case  $\alpha$ , if in  $\alpha$  C obtains in that way, in  $\alpha$  one is in a position to know that C obtains.<sup>34</sup>

Luminosity\* ( $L^*$ ) avoids the oddities of DeRose's suggestion, while it retains the crucial ability to block Williamson's argument. With  $L^*$  the argument ends at (5<sub>i</sub>), since  $L^*$  does not warrant the repeated application required for the *reductio* (Reed 2006, p. 307). The maximally specific way of obtainings Reed calls *central cases*, so the proposal is that we are always in a position to know that central cases of states of *feeling cold* or *suffering pain* obtain.

At this point in the dialectics we reach a mere clash of intuitions, since  $L^*$  seems to me wrong as a mark of the mental. Consider a case where high-caste Hi loves low-caste Lo, but where Hi is prevented from acknowledging her love to Lo due to a robust and severe internalisation of religious and cultural doctrines. Hi may never acknowledge her love for Lo, and she may not be capable on her own to break through such a set of robust internalised doctrines. This holds for any variants of maximally specified loving feelings for Lo, and it seems naïve and a piece of romanticism to say that, "there is a maximally specified way in which Hi could love Lo so that Hi's love for Lo would necessarily overcome all obstacles, which now prevents Hi from acknowledging her love for Lo". Perhaps Hi does not love Lo in that "special way"? But this would all be very awkward. Love is paradigmatically mental, so the failure of accommodation shows that  $L^*$  fails as an adequate mark of the mental.<sup>35</sup> My objection to  $L^*$  is that it does not hold universally. Like Davidson, I think there are cases where we can only grasp our own thoughts and achieve access to one's own mind by noticing what one says or does or being told so by others (1987, p. 15). It is a small step from this Davidsonian intuition to say that there are cases where the individual *cannot* notice what her mental states are. My intuition remains even if those cases involve affection, love and desire.

---

<sup>34</sup> See Reed (2006, p. 306).

<sup>35</sup> We find plenty of similar cases by turning to the realm of emotions, like *love*, *fear*, *grief* and so on. Peter Goldie has argued perspicuously that in many cases agents can have those emotions without noticing that they have them before later or upon being told so by others (2000). For example, consider the driver who in a near-accident discovers after the event how afraid he has been; or the embarrassed student who is so preoccupied with doing the right thing when walking into a room of famous scholars that he does not acknowledge his own embarrassment (evident by red cheeks etc) (Goldie 2000, p. 62). Goldie takes it that emotions can be had without a feeling, and I think we could easily replace "feeling" with "luminosity" in some of these cases.

### 2.3 A problem about belief

Evidently, if S *knows* A, that entails both *believing A truly* and *believing A*. Curiously, *believes truly* fails to be a mental state while the other two are mental, so it seems that there need not be a mental state  $M_c$  (believing truly  $p$ ) such that, necessarily, one is in  $M_c$  if and only if one is in  $M$  (believing  $p$ ) and C ( $p$  is true) obtains. However, there may still be a mental state  $M_k$  (knowing  $p$ ) such that, necessarily, one is in  $M_k$  only if one is in  $M$  and C obtains. We might ask whether it does not follow from the fact that knowing  $p$  entails believing  $p$  that the latter is a factor of the former? In other words, what is the proper relation between knowledge and belief? This section will discuss that issue.

In order to answer that question, we begin by considering the following entailment:

(Belief ( $B^*$ ))    If S knows that A, then S believes that A.

In the absence of any counterexamples,  $B^*$  is taken to be valid and articulates the thought that believing  $p$  is a necessary consequence of knowing it (Williamson 2000a, p. 42). If knowing is a mental state, the point is that the state of believing obtains necessarily if the state of knowing obtains. The problem is that *if*  $B^*$  necessitates belief as a *factor* of knowledge, then belief is among the necessary and sufficient conditions for knowing. But Williamson's working hypothesis requires that there are no such factors besides knowledge itself. Arguably, we can retain the validity of  $B^*$  by articulating anew the relationship between belief and knowledge so that it avoids factorising the latter.

Far from being a full-blown analysis, Williamson's strategy is to provide a "looser connection" between believing and knowing. A first approximation could be that believing is to have an attitude that is indiscriminable from knowing in that one believes  $p$  if and only if one has this indiscriminable attitude (Williamson 2000a, p. 46). Such a conception would validate  $B^*$  since one takes oneself to know  $p$  when one believes it. However, it would seriously misconceive the notion since it is consistent to claim that, "I believe  $p$ , even though I don't know that". Our clue to the notion of 'belief' should rather be that "to believe is to treat  $p$  as if one knew  $p$ " (2000a, p. 46). That would explain the intuitive indiscriminability, but it could also illuminate other features of believing.

An important characteristic of the factive attitudes (FAs) is that they have an essential reliability as premises in practical reasoning. Arguably, that reliance is also crucial for beliefs, but it can only be secured when knowledge sets the standard for appropriate beliefs and when

beliefs *aim at knowledge*.<sup>36</sup> The reliable beliefs are the ones that could have been known in the sense that they satisfy the aforementioned reliability-constraint.<sup>37</sup> The further one is from knowing  $p$ , the less appropriate it is to believe it. As a consequence, knowledge functions as a norm of beliefs in the sense that to know  $p$  is the best way to believe  $p$ . This is not to identify these mental states or make the one a factor of the other. The important lesson for us is to distinguish the mental state  $B(p)$  from  $K(p)$ , and say that  $S$  is *required* to be in a condition  $K(p)$  just because she treats  $p$  as if she knew it when she is in a condition  $B(p)$ . Belief without knowledge is a kind of “botched” knowing (Williamson 2000a, p. 47).

Intuitively that conception can be supported: Suppose  $S$  undertakes the following piece of practical reasoning, where she thinks that, “My lottery ticket is a loser, so if I keep it I will get nothing. However, if I sell it for a penny I’ll get a penny. Consequently, I’d better sell the ticket” (Hawthorne 2004, p. 29). It is clear that this reasoning is unacceptable, but the question is why? According to John Hawthorne, people tend to find the practical reasoning unacceptable because  $S$  does not know that her lottery ticket is a loser, which is in accordance with the above requirement on belief (2004, p. 30). So only knowledge warrants those beliefs we rely upon in practical reasoning.

Zeno Vendler has argued that grammatical divergences between beliefs and FAs indicate that they relate to different objects (Williamson 2000a, p. 43). If that is correct, then knowledge cannot be the best kind of believing. Consider the following sentences:

- (1) Pierre is aware of the *fact* that Penny is walking toward him.
- (2) Pierre is aware of the *true proposition* that Penny is walking toward him.

Evidently (1) is grammatically different from (2), but notice that if we replace “is aware of” with “believes” the difference will disappear. Vendler takes this to support that FAs relate to facts and not to true propositions, while beliefs relate to propositions. On the assumption that facts are something over and above true propositions, knowing is taken not to be a propositional attitude (PA) but only an FA. Since beliefs are PAs, knowledge cannot be the best kind of belief, which will jeopardise our above articulation of belief.

There is reason to be suspicious about Vendler’s claim. For the sake of argument, I shall put my own ontological qualms about the existence of facts (over and above true propositions) aside; we can still make some objections. For one thing, his claim invalidates  $B^*$  and is thereby in tension with standard views in the theory of knowledge (Harman 2002, p.

---

<sup>36</sup> Hawthorne has argued convincingly for the essential and intuitive link between knowledge and practical reasoning (2004, p. 29-31).

<sup>37</sup> See the previous section, and also chapter 3.

423). On  $B^*$ , knowing  $p$  entails that one believes *it*, but there is no obvious reason to expect that if S knows  $p$  S would also believe  $p^*$  when  $p \neq p^*$  (Williamson 2000a, p. 43). This is a cost on Vendler's proposal, since  $B^*$ , in lack of any clear case of a counterexample, is taken to be valid. Another outlandish implication is that if one utters a sentence "S knows that A" when A is false then A fails to refer (since there cannot be false facts). But then *no* sentence of the form "S knows that A" or "S does not know that A" can express propositions, which is absurd (2000a, p. 43). A third objection comes from surface grammar; consider the following:

- (3) Long before I knew those things about you I believed *them*.
- (4) I always believed that you were a good friend; now I know *it*.

On the assumption that the demonstratives "then" and "it" can be taken to *refer* to what is introduced earlier in these sentences, (3) and (4) support the claim that one believes and knows the same kind of object (Harman 2002, p. 424; Williamson 2000a, p. 43). Harman replies that the data is not conclusive, since the "it" in (4) can be taken to refer to the fact "that you were a good friend" rather than the true proposition that is believed. What happens is that the speaker uses "it" in a lazy manner and not referentially (Harman 2002, p. 424). In (3) Williamson takes "those things" to refer to propositions, but this seems odd since a replacement of "those things" with propositions will sound queer (Harman 2002, p. 425).

I will grant Harman his point, but he abstains from answering the graver objection; i.e. that instances of an utterance "S knows that A" will fail to refer when A is a falsehood. Williamson suggests a plausible answer to this objection: Simply interpret "that A" as an elliptical phrase for "the fact that A", and then analyse this expression in accordance with a standard Russellian theory of definite expressions (2000a, p. 43).<sup>38</sup> That move will avoid the objection from lack of reference, since a reference for "the fact that A" can be replaced by a function  $f$  on the proposition  $p$  that is expressed by A. Thus to know that A is to know  $f(p)$ , which is to stand in a complex relation expressed by "know", "the" and " $f$ " to " $p$ " (Williamson 2000a, p. 43). If we can use the word "know" to express that complex relation to a proposition, then any talk of facts (over and above true propositions) will evaporate. All in all, I take it that Vendler is wrong, and that knowledge relates to true propositions, which makes our above conception of belief as a knowledge-governed mental state tenable.

---

<sup>38</sup> See Russell (1905; 1919a). The basic idea is to analyze a definite description ("the so-and-so") into a complex quantified expression. For instance will the famous bald French king be analyzed as:  $(\exists x) (Kx \ \& \ Bx \ \& \ (y) (Ky \supset x = y))$ . The important point is that a referent of a French king is no longer required in the quantified expression. If that can be done with false facts too, then talk about false facts will not require reference to false facts, which is good since false facts cannot exist.

## 2.4 Twin Earth

According to *the internalist conception of mind* there is a sharp delineation between the mind and world external to the agent, and the individual's mental characteristics are completely constituted by what goes on within the skin (or skull) of the individual (McCulloch 1995, p. xii; Williamson 2000a, p. 49).<sup>39</sup> A consequence of this claim is that knowledge cannot be purely mental, since knowledge is a complex function of mind and world variables. This internalist metaphysical constellation of mind-states takes knowledge to be a composite hybrid "crying out for analysis" into its internal and external components (Williamson 2000a, p. 5). Recent progress in the philosophy of mind has put an alternative conception out in logical space, which takes mind and world as dependent variables (2000a, p. 5). This is *the externalist conception of mind* on which mind is inseparable from the body in which it subsists and the surroundings in which that body lives (McCulloch 1995, p. xii).<sup>40</sup> A natural move for Williamson is to embrace the latter and dispute the former, in order to support his claim that knowledge is a mental state. In this section I will provide the structure of Williamson's *radical externalism*, before we confront internalism and the objections to Williamson in the next sections. First, let us become familiar with Williamson's jargon.

We can identify the *internal* as the total internal physical state of an agent S at some time *t*, and the *external* as the total physical state of the environment external to S at *t*. We also define a *case* as a possible total state of a system consisting of S at *t* paired with an external environment. A *condition* is something that obtains (or fails to obtain) in each case, and a condition C *entails* D if and only if for every case  $\alpha$  if C obtains in  $\alpha$  so does D. Two conditions C and D are *identical* if and only if for every case  $\alpha$  C obtains if and only if D obtains. A case  $\alpha$  is *internally alike*  $\beta$  if and only if the total internal physical state of S in  $\alpha$  is identical with the total internal physical state of S in  $\beta$  (2000a, p. 51-2). They are *externally alike* if and only if the total physical state of the external environment is exactly the same in  $\alpha$  and  $\beta$  (2000a, p. 66). Call a condition C *narrow* if and only if, for all internally alike cases  $\alpha$  and  $\beta$ , C obtains in  $\alpha$  if and only if it obtains in  $\beta$ . A condition is *broad* if and only if it is not narrow (2000a, p. 52). Now we can say that a *state*<sup>41</sup> S\* is narrow if and only if whether S is in S\* at *t* depends only on the internal qualitative state of S at *t*. Similarly, we can define a

---

<sup>39</sup> McCulloch characterizes internalism as the position that draws a sharp delineation between mind and world exterior to the agent. 'Exteriority' is an ontological concept that concerns the dependence of the existence of one thing on another. Roughly speaking, X is exterior to Y iff Y could in principle exist just as it were whether or not X exists (McCulloch 1995, p. 15). Thus McCulloch's characterization of internalism entails a self-containment claim, where the states of the mind could exist as they are without an exterior world, which makes mental content independent of the external world (1995, p. 184). We will mainly discuss a weaker internalism.

<sup>40</sup> In this sense externalism is the negation of internalism (Williamson 2000a, p. 52).

<sup>41</sup> A state is a condition that an agent can be in (Williamson 2000a, p. 52).

*broad state* as those that are *not* narrow (2006, p. 291). Finally, a state is *environmental* if and only if it depends only on the total state external to S at *t* (2005a, p. 66; 2006, p. 292).

If two internal duplicates are situated in two different cases they are internally alike and necessarily share identical narrow conditions. They may still differ in their broad conditions depending on whether the environmental states are different (i.e. whether the cases are not externally alike). We can define *internalism* as the claim that all *purely mental* states are narrow, and *externalism* as the denial of that claim (2000a, p. 52). Then internalism will predict that two internal duplicates share all their purely mental states even though they are situated in externally different cases. Our definition may also elucidate the internalist resistance to allow knowledge as a purely mental state. For suppose that A and B are internally alike and thus share all narrow conditions. Then they can be situated in different cases  $\alpha$  and  $\beta$ , where A knows *p* in  $\alpha$  while B does not know *p* in  $\beta$  due to external differences.<sup>42</sup> Since that is possible, knowledge cannot be determined only by narrow conditions and thus disqualifies as mental (2000a, p. 54-6).<sup>43</sup>

Williamson seeks to dispute this consequence, and in order to do that he seeks support in the externalist conception of mind. In a series of papers, Hilary Putnam (1973; 1975; 1981) argued that linguistic content depends constitutively on environmental factors that are *external* to the agent. According to Williamson, that claim makes content-ascriptions through propositional attitudes (PAs) a source of broad conditions in natural language, and he takes this to have consequences for the philosophy of mind (2000a, p. 53). Since we specify thoughts, beliefs, desires and intentions in terms of subordinate sentential clauses (that-clauses) of propositional attitudes, it is thought that those states themselves can be shown to be broad. It is standard to take the content-clauses of PAs to have *oblique occurrences* in the sense that they are not freely exchangeable or intersubstitutable *salva veritate* with extensionally equivalent expressions (Burge 1979, p. 599). Let me illustrate: We identify water with H<sub>2</sub>O. Suppose Pierre believes that water is refreshing and yet fail to believe that H<sub>2</sub>O is drinkable at all. If we say, “Pierre believes that water is refreshing”, we cannot just substitute the occurrence of ‘water’ with its extensional equivalence ‘H<sub>2</sub>O’ without altering its truth-value. Pierre would probably oppose and say, “Hey, I don’t believe that H<sub>2</sub>O is drinkable at all!”<sup>44</sup> Since PAs function as a primary means for identifying a person’s

---

<sup>42</sup> E.g. in  $\alpha$  *p* is true and in  $\beta$  *p* is false.

<sup>43</sup> I omit the details of this argument, but see Williamson (2000a, p. 54-6).

<sup>44</sup> Due to this obliqueness propositional attitudes are often taken to be essential constituents in so-called *intentional discourse* (*oratio obliqua*), but these semantic interpretations are highly controversial. The phenomenon of apparent obliqueness is the main source of an extensive bulk of literature in the philosophy of language and constitutes one of the fundamental problems in the field. Frege famously took the phenomenon to necessitate that we invoke a notion of *sense* in addition to *nominatum* (reference). Frege took the reference of an oblique

intentional mental states, and these identifications are just expressions used in content-clauses, we may derive a radical form of externalism from Putnam's argument about externalist contents (Burge 1979, p. 599). Let us argue the case in more detail. Consider the following:

TWIN EARTH:<sup>45</sup> Suppose that Pierre believes *that water is refreshing*, and that he ascribes this thought to himself by uttering, "I believe that water is refreshing". Abbreviate that content *W*. It is then natural to say that Pierre has ascribed *W* to himself. Suppose further that at Twin-Earth Pierre has an atom-by-atom physical duplicate that is his internal *doppelgänger*. Call him Twin-Pierre. The two Pierres are internally alike, and thus share all narrow conditions. However, there are external differences since on Twin-Earth, even though it is as close to Earth as can be, there is no water (H<sub>2</sub>O). On Twin-Earth there is only water<sub>TE</sub> (XYZ). Water<sub>TE</sub> is a substance that in all appearance is similar to ordinary water (it is not coloured, it is drinkable, it freezes at 0°C, and the Twin-Earthians call it 'water'), but which nevertheless *is* different than water due to its different molecular structure.<sup>46</sup>

The question is whether Twin-Pierre would attribute *W* to himself by saying, "I believe that water is refreshing"? But why should we take Twin-Earthian lingo to be speaking about water (H<sub>2</sub>O), when all the time the Twin-Earthians relate to water<sub>TE</sub> (XYZ)? Putnam's intuition is that only Pierre attributes *W* to himself, while Twin-Pierre would attribute some other content to himself, *W\**, if he were to utter the same words as Pierre (1975, p. 585). The result is that two internal duplicates have identical internal physical states and narrow conditions, and yet they attach different contents to their words. Since the only difference between the two cases is environmental and external to the agents, it shows that content must be broadly determined and hence that content depends constitutively on external conditions.<sup>47</sup>

So far we have discussed content-externalism, but for this to do any explanatory work for Williamson we need to prove that there is also an essential broadness to attitudes. The

---

that-clause to be the sense of the utterance (1949). Philosophers allergic to these intentional entities have sought other solutions to the problem that-clauses in oblique occurrences in order to recover our pre-Fregean semantic innocence. Davidson worked out a solution where the 'that' in the that-clause is taken demonstratively pointing to the clause or sentence that immediately follows it, and where the latter is mentioned or displayed and not used (1968, p. 368-9). The phenomenon is especially challenging to adherers of the direct reference semantics for proper names (i.e. where the meaning of a name is simply its referent). This is demonstrated in Kripke's case of puzzling Pierre, who believes that "London is ugly" at the same time as he thinks that "*Londres est jolie*" (1979 p. 414-5); how could this be if the meaning of the name just is its referent? Lately the obliqueness has been questioned by Scott Soames (2002, chapter 6).

<sup>45</sup> This is merely a brief paraphrase of Putnam's famous Twin-Earth argument for externalist semantics. See Putnam (1973) for the original phrasing of this argument, and also his (1975) for an extended version of the same argument. In a later paper (1981) he also attempts to solve the epistemological problem of scepticism by appealing to externalist semantics and the argument from Twin-Earth.

<sup>46</sup> Water<sub>TE</sub> has molecular structure XYZ while water has H<sub>2</sub>O. McCulloch exploits Locke's terminology and says that the *real essence* of these two substances is different, while their *nominal essence* or *stereotype* is identical (1995 p. 161-3). According to Locke, the nominal essence of water would be how we define it or the idea we connect with it, while its real essence is its usually unknown constitution ("the being of anything whereby it is what it is") (Locke 1998, p. 333).

<sup>47</sup> Note that the two Pierres may not be aware of these contents: If Earthian lingo tracks H<sub>2</sub>O, then before scientists discovered this fact about water our word tracked something that was unknown to the community of speakers. Putnam's semantic hypothesis is that the meaning of words is partly determined by causal- and/or social relations. In other words "'meanings' just ain't in the head!" (1975, p. 587). To speak with John Locke, we could say that Putnam's intuition is that 'water' tracks real essence. Locke himself took substance-words to track only nominal essence, since real essence was largely unknowable. Putnam's view can therefore be phrased as "the doctrine that understanding tracks real essence". See McCulloch (1995, p. 162-3).



following *reductio* might be sufficient.<sup>48</sup> Suppose that beliefs are narrow and that the two Pierres hold a glass of H<sub>2</sub>O (5) and XYZ (6) respectively (Williamson 2006, p. 293-4):

- (5) Pierre is holding a glass of water.
- (6) Twin-Pierre is not holding a glass of water.

On the assumption that beliefs are narrow, Pierre will express (7) while Twin-Pierre will express (8) if they both were uttering, “I am holding a glass of water”:

- (7) Pierre believes that he (Pierre) is holding a glass of water.
- (8) Twin-Pierre believes that he (Twin-Pierre) is holding a glass of water.

By assuming that the schemas (T) and (F) govern true and false beliefs, we get (9) as an instance of (T) and (10) as an instance of (F) (2006, p. 294):

- (T) If S believes that  $p$ , S believes truly that  $p$  if and only if  $p$ .
- (F) If S believes that  $p$ , S believes falsely that  $p$  if and only if not  $p$ .<sup>49</sup>
- (9) If Pierre believes that he [P] is holding a glass of water, Pierre believes truly that he [P] is holding a glass of water if and only if he [P] is holding a glass of water.
- (10) If Twin-Pierre believes that he [T-P] is holding a glass of water, Twin-Pierre believes falsely that he [T-P] is holding a glass of water if and only if he [T-P] is not holding a glass of water.

But from (5), (7) and (9) we can conclude (11); and by the same token we can reach (12) by (6), (8) and (10) (Williamson 2006, p. 294):

- (11) Pierre believes truly that he [P] is holding a glass of water.
- (12) Twin-Pierre believes falsely that he [T-P] is holding a glass of water.

It would be extremely awkward to attribute false beliefs on Twin-Pierre and not on Pierre, and this absurdity forces us to retract the assumption. What this shows is that the narrow conditions are not sufficient to determine attitudes, since, if they were, we should have no trouble attributing the same belief to both Pierres, and hence attribute a falsehood to Twin-Pierre. Since those attitudes were left *undetermined* by all internal qualities, we can conclude that beliefs are *not* narrow conditions; and so by definition they are *broad* (2006, p. 294-5). Note that this argument has an extremely wide application, since *most* of our content-

---

<sup>48</sup> Williamson presents this argument at two places (2000a, p. 53; 2006, p. 293-5). In what follows, I combine aspects of both presentations for the sake of expository convenience.

<sup>49</sup> (T) and (F) are elucidated in more detail in chapter 6. See also Williamson for a phrasing in terms of utterances (1996, p. 188).

ascriptions in natural language attribute broad rather than narrow conditions (Burge 1979, p. 601; Williamson 2000a, p. 53).<sup>50</sup> The question is therefore whether one can reconcile these results with a conception that takes purely mental states to be narrow only.

## 2.5 Factorisation and the internalist divide and conquer

As a consequence of the Twin-Earth thought-experiment, we can take attitudes to be broad conditions that constitutively depend on external factors. This puts pressure on the internalist, since her criterion of the purely mental is that those states depend only on narrow conditions and the internal physical states of the agent. It would be very awkward if the internalist had to say that broad attitude-ascriptions did not attribute mental content, so in response she typically applies a *divide-and-conquer* strategy, which attempts to *factorise* the problematic broad attitudes into narrow and broad components (Williamson 2000a, p. 53). The internalist thinks that she can concede that content- and attitude-ascriptions in natural language express broad rather than narrow conditions, while at the same time she can insist that those ascriptions fail to reflect the structure of the underlying facts (2000a, p. 53). That puts the burden of proof on the internalist, since she must be able to show a narrow non-factive attitude to any broad factive attitude. If successful, she accommodates the letter of Putnam's argument while retaining the internalist spirit. Let us attend to her efforts.

From a psychological perspective, broad conditions are often thought to be beside the point, and a common thought is that belief is the purely mental component of knowledge and thus what is relevant from a psychological perspective (2000a, p. 49). Hence a natural option for the internalist is to factorise knowledge in terms of belief. In the following (1<sub>i</sub>) captures the fundamental internalist claim, while (2<sub>i</sub>) is the internalist commitment to factorisation:

- (1<sub>i</sub>) For all cases  $\alpha$  and  $\beta$ , if  $\alpha$  is internally like  $\beta$ , then one is in exactly the same mental state in both.
- (2<sub>i</sub>) For all propositions  $p$  and cases  $\alpha$ , if in  $\alpha$  one *believes*  $p$  then in some case  $\beta$  one is in exactly the same mental states as in  $\alpha$  and one knows  $p$ .<sup>51</sup>

On (2<sub>i</sub>), the purely mental contribution of a knowledge-relation is belief, and hence knowledge is factorised into belief plus some broad X-factor. Importantly, (2<sub>i</sub>) invalidates Williamson's claim that knowledge is a purely mental state. But there is reason to believe that (2<sub>i</sub>) is false: For suppose one believes a necessary falsehood like "12 + 5 = 13". That belief cannot be known since it is false in any case, which constitutes a counterexample to (2<sub>i</sub>).

---

<sup>50</sup> We can create an analogous argument whenever people have less than complete understanding of a thought or notion (Burge 1979, p. 601).

<sup>51</sup> See Williamson (2000a, p. 55-6).

Can we reformulate it? A suggestion is that we replace “believes  $p$ ” in (2<sub>i</sub>) with “believes  $p$  truly” (Williamson 2000a, p. 57). The problem is that *believing truly* is a broad condition and thus fails to provide a factorisation. A further problem is the fact that one can come to believe propositions on irrational grounds that stand in the way of knowledge. This suggests that belief is too unspecific to constitute the mental component of knowledge, since it does not sufficiently exclude what knowledge-states leave out (2000a, p. 57).

Another suggestion is therefore that we replace “believes  $p$ ” with “rationally believes  $p$ ” (2000a, p. 57). But suppose I try to denote an individual tree by a demonstrative use of an expression, where I say, “That tree is green; therefore, a tree is green”.<sup>52</sup> Suppose further that my denotation fails to pick out a tree since I suffered from a hallucination. Then I fail to express a proposition by my demonstrative use of “that tree”, and as a result my mental state have a corresponding lack of singular belief. My inference will therefore lack a singular belief to serve as a premise, and either: (a) the belief “a tree is green” is rational; or (b) the belief is not rational. If (a), then according to our principle someone can be internally identical to my state and come to know that “a tree is green”. But that is odd, because then someone could establish knowledge by using a premise in her inference that lacks a singular belief. And on (b), I fail to form a rational belief due to external factors, which makes rational belief dependent on broad conditions. Thus neither option is of much use to the internalist (Williamson 2000a, p. 58). Another thing that goes against the principle is that considerations of rational belief depend on considerations about knowledge. For consider a case  $\alpha$  where A believes that, “Ticket #66 will not win the lottery” ( $T$ ), on the basis of a high probability less than 1. If in case  $\beta$  B is in the same mental state as A in  $\alpha$ , and B believes  $T$  on same probabilistic grounds, then one does not know  $T$  in any of these cases (after all, if one knew one would not have bought it). But then it follows from our principle that one does not rationally believe  $T$  in  $\alpha$ . Again, a factorisation-effort has failed.

So far our discussion has reached inconclusive support for the failure of factorisation. However, we can make a stronger argument through considerations on *primeness*. According to Williamson, cognitive states that explain action at an appropriate level of generality are broad in a way that makes their factorisation impossible in principle (2006, p. 292). We lose the explanatory virtue of *significant generalisation* available at a higher level by moving down the level of basic actions, so that the internal conditions cannot play the distinctive role in explanation that internalism predicts (2000a, p. 65, 82; 2006, p. 296). On the assumption

---

<sup>52</sup> Here the “that” is the expression demonstratively used. This is in accordance with Kaplan’s *Dthat*-operator, which names the demonstrative when it is demonstratively used. *Dthat*-operator employs the sense of the expression to fix the denotation (i.e. the demonstratum), but once that is fixed the sense of the operator no longer partakes in the corresponding proposition (Kaplan 1978, p. 332).

that psychological intentional states are relevant for explaining action, *prime* conditions are required since those states tend to be prime when they are individuated by natural language predicates (2006, p. 300). I will first briefly provide Williamson's argument before I make a critical scrutiny of the alleged necessity of prime conditions.

Let us avail ourselves of some terminology: Define a condition *C* as *composite* if and only if it is a conjunction of a narrow condition *D* with an environmental condition *E*, and call a condition *prime* if and only if it is not composite (2000a, p. 66; 2006, p. 298). Note that all narrow conditions must be trivially composite and thus that all prime conditions are broad, even though none of these relations go vice versa (2006, p. 298). We can now discern *C* as prime by exhibiting three cases,  $\alpha$ ,  $\beta$  and  $\gamma$ , where  $\gamma$  is internally like  $\alpha$  and externally like  $\beta$ , and where *C* obtains in the first two cases without obtaining in the last (2005a, p. 68).<sup>53</sup> Conversely, *C* is composite if no such triple of cases exists (2000a, p. 68). Recall that on internalism there is a tendency to think of narrow mental conditions as underlying the broad conditions and doing all the causal work needed to explain action. This conception renders all mental conditions trivially composite due to their alleged narrowness. The internalist thinks that she can reduce any mental condition to a composite *C*, and isolate a condition, *virtual-C*, which obtains in a case  $\alpha$  if and only if *C* obtains in some case internally like  $\alpha$ , as the purely mental reality underlying *C* (2000a, p. 66-7). For knowledge the candidate *virtual-C* is belief. The composite also contains an environmental condition, *outward-C*, which stands to the external as *virtual-C* stands to the internal (2000a, p. 67). What we need to show is that most mental conditions are prime, so let us do that by employing triples of cases that show that *C* can fail to obtain when both *virtual-C* and *outward-C* obtain (2000a, p. 67).

Arguably *to see water* is a mental condition. Now let  $\alpha$  be a case where one sees water normally with one's right eye, while one's left eye receives light rays that are like those it would receive from water, while in fact they are emitted by a waterless device just in front of the eye. Suppose that a head injury prevents one from further processing of input from one's left eye. One obviously sees water in this case. Let  $\beta$  be a case that differs from  $\alpha$  by reversing the roles of the two eyes: In  $\beta$  one sees water with the left eye, while the right eye receives the light rays. Again, one suffers from a head injury, so the processing from one's right eye is blocked. One sees water in  $\beta$  too. Let  $\gamma$  be a case that is internally like  $\alpha$  and externally like  $\beta$ . Then  $\gamma$  is a case where a head injury prevents further processing of input

---

<sup>53</sup> I omit the proof for this test of primeness, but see Williamson (2006, p. 305n7).

from one's left eye (as in  $\alpha$ ), and where one's right eye does not receive light rays from water (as in  $\beta$ ). One does not see water in  $\gamma$ . Hence that condition is prime (2000a, p. 69-70).

We can argue for the existence of an analogous triple of cases with *epistemic conditions*: Let there be a case  $\alpha$  where one *knows by testimony* that an election was rigged ( $E$ ). In  $\alpha$  A is trustworthy and one trusts her, while B is unreliable and not to be trusted. Both A and B tells you that  $E$ . Let  $\beta$  be a case that differs from  $\alpha$  by reversing the roles of A and B. Again, one knows by testimony that  $E$  since A tells you that  $E$  and B tells you the same while you trust only the latter (which is good since now B is trustworthy). Consider the case  $\gamma$  that is internally like  $\alpha$  and externally like  $\beta$ . In  $\gamma$  one does not trust B because one does not trust B in  $\alpha$ , and  $\gamma$  is taken to be internally like  $\alpha$ . Furthermore, in  $\gamma$  A is not trustworthy because he is not reliable in  $\beta$  and  $\gamma$  is externally like  $\beta$ . Thus  $\gamma$  is a case where neither A nor B is both trustworthy and trusted, and that destroys knowledge in  $\gamma$ , since only testimony warrants knowledge in these cases. Consequently, knowing by testimony is prime (2000a, p. 72).

Importantly, these results can be delivered with any triple of cases that involves object-dependent contents of propositional attitudes. Hence the argument for primeness generalises, and its result is just as applicable as the outcome of the general argument for externalism (Williamson 2000a, p. 71). If successful it generates strong support for externalism at the same time as it refutes internalism. Let us see if it survives critical scrutiny.

According to Jonathan Cohen and Paul D. Magnus, Williamson's conclusion is in need of important qualification, since his argument can only refute a specific strand of internalism, which leaves the most plausible version untouched. Recall that internalism factorised broad attitudes and analysed "S believes that  $p$ " as saying: (a) S stands in some computationally specified relation to a mental representation  $\phi$ ; and (b)  $\phi$  means that  $p$  (Cohen and Magnus 2003, p. 39). The internalist accommodated the argument from Twin Earth by acknowledging the broadness of (b), while retaining (a) as narrow. Note also that Williamson's argument assumes that case  $\gamma$  is possible only if one can combine conditions from  $\alpha$  and  $\beta$  that are independently narrow (NC) and broad (BC) (2003, p. 42). In other words, Williamson assumes that internalism *must* endorse the *principle of free recombination* in which NC and BC are independent variables (Williamson 2000a, p. 73).

Cohen and Magnus think that this principle misreads condition (a) and (b) as if they were independent, while the most plausible interpretation will read them as a quantified statement bound by a single quantifier that ranges over both conditions. Then internalism says this:  $B(s,p) \equiv \exists \phi (R(s,\phi) \ \& \ M(\phi,p))$  (Cohen and Magnus 2003, p. 42). Here S believes  $p$  if

and only if two conditions obtain *when* bound by a single quantifier. Hence we cannot break up (a) and (b) and freely re-combine these conditions, which is required for Williamson's argument to work (2003, p. 43). The argument shows that no *reasonable* internalism can conceive attitudes as conjunctions of narrow- and broad states, where those states are independent and can obtain independently of one another. But an internalism that factorises without asserting this independence is untouched by this argument. Hence it seems that Williamson's argument can only take down one version of internalism.

## 2.6 The causal nexus

According to Cohen and Magnus, cognitive science and special sciences more generally conceive valid explanations as subsuming individuals under causal generalizations (2003, p. 38). Since mental states are invoked in order to explain human action, they are required to be *causally efficacious* in a relevant manner (2003, p. 38). The question is whether broad states and prime conditions are relevant in this causal nexus in a sense that give them any explanatory value. Williamson takes it that they are causally involved in the relevant manner, and that they have *ineliminable* explanatory *value*. This section attends these issues.

According to Cohen and Magnus, it is often thought that the relevant psychological causal relations must be among narrow conditions, because the workings of a causal nexus are local (2003, p. 38). Admittedly, talk about "local causality" is somewhat metaphorical, but I think Cohen and Magnus take it to be something along the following lines: Whether S does  $\phi$  or  $\varphi$  depends on which causal chain S initiates, the choice of which depends only on the internal characteristics of S. Causality is local in the sense that the choice of causal chain (and hence action) must always be mediated through S's internal characteristics, and hence they think that those characteristics are sufficient to explain why S does  $\phi$  rather than  $\varphi$ . Importantly, that renders knowledge causally and explanatorily irrelevant (Cohen and Magnus 2003, p. 38; Williamson 2000a, p. 60-1). Williamson admits that we have reason to be suspicious of purported mental states if they do not appear *non-reductively* in psychological explanations; however, the explanatory redundancy of knowledge (and prime states) must be cashed out first, and Williamson does not think that we can do that (2000a, p. 61). A scientist, who restricts herself to narrow- and environmental conditions, cannot sufficiently explain the behaviour of agents, since many explanations of *why* S did  $\varphi$  are not sufficient without appeals to prime conditions (Harman 2002, p. 417).

Cognition is supposed to help agents achieve their goals by enabling more successful interaction with their environment, so it is important that one's narrow states are in line with the environmental states (Williamson 2006, p. 14). For instance, will a series of internal states I1, I2, I3 etc with the right combination of environmental states E1, E2, E3 etc yield the state of *seeing water*.<sup>54</sup> Our previous section showed that *seeing water* is prime, and that folk-psychological intentional states individuated by use of natural language predicates tend to be prime. Williamson's point is that we cannot reduce them to composites of narrow conditions conjunctively related to external conditions, without loss of explanatory value (2006, p. 19). Just consider a case where A is the condition that S performs  $\phi$ , and  $\Phi$  is the appropriate circumstances for A to obtain. Let B be the condition that A believes  $\Phi$  obtains, and K the condition that A knows  $\Phi$  obtains. Now define B\* as a condition that S *believes truly* that  $\Phi$  while she does not *know* that  $\Phi$  (B\* =<sub>df</sub> B & f & ~K) (Cohen and Magnus 2003, p. 45-6). Williamson's claim is that for some significant range of cases K there is a higher correlation with A than is the case with B\*. Hence the correlation between K and A is greater than between A and B\*, and so S's knowledge correlates higher with S's doings than mere beliefs do. As a result, knowledge is a better predictive factor of S' behaviour than true belief. Williamson's claim is that the following holds for a range of plausible cases:  $P(A|K) > P(A|B^*)$  (Cohen and Magnus, 2003, p. 46).<sup>55</sup>

Perhaps the following is such a case: Suppose you are home when B knocks on your door. You hope B will go away, so you consider whether or not to reply. Plausibly, your choice will depend on whether you think B concludes (wrongly) that you are not home or concludes (rightly) that you are hiding away (and thereby becomes cross with you). If B knows that you are home, this will be a better prediction whether B will take offence than if B merely believes (truly) the same. After all, a *knower* is less likely to revise her beliefs than a mere *true believer* (2000a, p. 86). A true believer may believe truly that C obtains in a case  $\alpha^*$  that is sufficiently similar to other cases  $\alpha_1, \dots, \alpha_n$ , where C does not obtain. Scrutiny may therefore reveal the true belief as unreliable, which may have consequences for her choice of action and lower the probability of doing  $\phi$  based on the belief that C obtains. Also, a true belief may be based on false beliefs. For instance, if B believes that you are at home on the basis of the false belief that your car is parked outside the house, and then B will probably lose the belief that you are at home upon being told that the car belongs to someone else. But

---

<sup>54</sup> E.g. looking ahead (I1) with the ocean in front (E1), or looking down (I2) while standing on a pier with the ocean beneath (E2), etc.

<sup>55</sup> Read  $P(x|y)$  as "the probability of x conditional on y".

that information would not change his belief if he knew that you were at home, since what is known could not be reliably based on this false belief (2000a, p. 86-7).

So is this a case where knowledge is an ineliminable feature of explanation? According to Cohen and Magnus, there is no need to accept the superior explanatory force of knowledge in this case, since that evaporates once we take B's belief to be of a kind that survives scrutiny (belief\*). In that case, B's belief\* will do as much explanatory work as knowledge since neither is liable to revision on scrutiny (Cohen and Magnus 2003, p. 47). But this move seems to be *ad hoc*, since I can know in advance whether B knows *p* but it seems impossible to discern ordinary true beliefs from true beliefs\* *before* scrutiny. As such, knowledge is a better predictor of scrutinizing behaviour. Consequently, Cohen and Magnus' notion has significantly less explanatory value in a case where B knocks on my door. A stronger objection to their notion is that it seems to me that the most plausible way we can carve out beliefs\* is by way of reliability and *epistemic safety*. As we have seen, true beliefs that satisfy the reliability-constraint can be known.<sup>56</sup> So it seems that Cohen and Magnus' notion just captures those beliefs that can be known, and hence the difference between them and Williamson seems to be no more than the use of a different nametag.

## 2.7 Conclusion

The above discussion only scratches at the surface of a vastly complex and important field of research in contemporary philosophy of mind and action, yet I think it is safe to conclude that Williamson has made a strong case for his account of knowledge as a mental state. The most important results seems to be that we may achieve reflective understanding of knowledge through the modest positive account of 'knowledge' (MPA) coupled with the general framework of an externalist conception of mind. We have also seen reasons to support the ineliminability of knowledge and prime conditions in explaining human action and behaviour, as well as reasons to believe that the internalist conception of mind is insufficient. Importantly, Williamson's argument failed as a general rejection of internalism, since it could not challenge internalists that do not conceive attitudes as conjunctions of narrow and broad conditions. The price for this kind of internalism is that narrow conditions must be closely conjoined with broad conditions. Importantly, Williamson acknowledges this restriction on his own argument, but he thinks that the argument for ineliminability of prime states in psychological explanation proves that this internalist variant cannot re-establish the hegemony of narrow conditions in psychological explanation (2000a, p. 90, 92). Knowledge *is* needed.

---

<sup>56</sup> See section 2.2 and the subsequent chapter.



## Chapter 3: Knowledge as justification

*Knowing the extent of our capacities will hinder us from useless curiosity, scepticism and idleness.*

John Locke

*If truth is a hit, knowledge is a safe hit.*

Timothy Williamson

A key feature of Williamson's conception of knowledge is that the notion of knowledge itself is a – and perhaps *the* – key explanatory notion in epistemology. In the previous chapters we have looked at the central elements of his account of the concept of knowledge; now we turn to issues concerning its explanatory role in epistemology. The issue of justification has always been a central theme in epistemology, but philosophers have mainly sought to illuminate and explain the notion of knowledge through the notion of justification, rather than the other way around. According to Williamson, that picture must be turned on its head. He suggests that we equate one's body of evidence with one's body of knowledge ( $E = K$ ), and that we conceive knowledge primarily as a justifier. In this sense, Williamson commits himself to a modest foundationalist conception of justification, where, again, knowledge comes first. I begin in section 3.1 by providing the basic structure of Williamson's equation ( $E = K$ ), and then discuss his main argument to support that equation. In section 3.2 I discuss some objections and replies to this equation, whereas in section 3.3 I take up some issues about evidence that are related to the anti-luminosity of mind.

### 3.1 Evidence

In chapter 1 we saw that Williamson is sceptical about the prospect of a successful analysis of knowledge, and that he suggests that our working hypothesis should be that knowledge is unanalysable. Importantly, he thinks that there is reason to believe that any standard of justification that is independently accounted for without employing the notion of knowledge is insufficient to immunize an analysis of knowledge from Gettier counterexamples. Hence any *immune* analysis of knowledge in terms of *justification* will risk vicious circularity (2000a, p. 4). Again, Williamson takes the long line of failed analyses to indicate strategic failure: It has been showed that *what is justified* need not be knowledge, but yet it need not follow that *what justifies* is not knowledge (2000a, p. 185). So perhaps it was wrong from the start to seek illumination of knowledge through justification? Perhaps it is more rewarding if we turn the direction of explanation, and seek to elucidate justification and evidence through knowledge? The following section will explore that option.

Let me start by fleshing out the position: According to Williamson, we may *equate* knowledge with *what justifies* beliefs, and thereby illuminate justification through knowledge. The state (or condition) you must be in, *in virtue of which* you are justified in believing  $p$ , I shall call a *justification-maker* or *justifier* (Pryor 2005, p. 182). A justification-maker is a state that makes it epistemically appropriate for  $S$  to believe  $p$ , so the suggestion is that the only appropriate justification-maker is the state of knowledge. If  $S$  believes  $p$ , then that belief is justified if and only if what  $S$  knows justify it (usually some set of propositions  $q_1, \dots, q_n$  other than  $p$ ) (2000a, p. 185). It is standard to speak of justifiers as one's evidence, so Williamson's claim is equivalent to the principle that only knowledge constitutes evidence, and hence that  $S$ 's evidence is equated with  $S$ 's knowledge:  $E = K$  (1996, p. 245; 2000a, p. 185).

An evident function of evidence ( $e$ ) is to speak in favour of a claim or hypothesis ( $h$ ); hence  $e$  should raise the probability of  $h$ . According to Williamson, a natural idea is that  $S$  has a *body of evidence* and that  $e$  is evidence only when it is included in that body of evidence. We can capture the claim that  $e$  speaks in favour of  $h$  by the following principle:

EV  $e$  is evidence for  $h$  for  $S$  iff  $S$ 's evidence includes  $e$  and  $P(h|e) > P(h)$ .<sup>57</sup>

On *EV*, something is evidence for a hypothesis when and only when it raises the probability for that hypothesis. Curiously, according to *EV*, some  $e$  is evidence for  $h$  only if  $e$  is evidence for itself. For  $P(e)$  can neither have the probability 0 (otherwise  $P(k|e)$  would be ill-defined), nor 1 (otherwise  $(P(k|e) = P(h))$ ). Observe that it would be well-defined to say that  $P(e|e)$  has value of 1 and is greater than  $P(e)$ . But then  $e$  satisfies *EV*, and thereby  $e$  is evidence for itself (Williamson 2000a, p. 187). It seems circular to say that  $e$  can only be evidence if it is evidence for itself, because that implies that one could take  $e$  to speak in favour of itself. But that circularity does not make it trivial whether  $S$ 's evidence includes  $e$ , since when one equates evidence with knowledge, that makes  $e$  a part of  $S$ 's body of evidence only if  $S$  knows  $e$ . To know  $e$  is evidently not a trivial matter (2000a, p. 187).<sup>58</sup> Yet, there is a sense in which it is conversationally inappropriate to cite  $e$  in support of itself when  $e$  is challenged. It is a matter of discursive propriety to temporarily withdraw propositions under challenge when their status as evidence has been challenged (2000a, p. 187-8).<sup>59</sup>

It is not obvious that  $E = K$ , so Williamson's suggestion requires support. He thinks that he can get it from the following argument:

<sup>57</sup> See Williamson 2000a, p. 187. Read  $P(x|y)$  as "the probability of  $x$  on  $y$ ".

<sup>58</sup> It is not trivial whether  $S$  knows  $p$ , but notice that *if*  $S$  knows  $p$  then  $p$  has trivially a probability of 1 on its evidence.

<sup>59</sup> In chapter 6 I show how sceptics dialectically exploit such features of discourse to undermine knowledge (Williamson 2000a, p. 188).

- (Ei) All evidence is propositional.
- (Eii) All propositional evidence is knowledge.
- (Eiii) All knowledge is evidence.
- (K) ∴ All and only knowledge is evidence.<sup>60</sup>

The argument is valid, so the question is whether its premises are true. Let us begin with the first premise: Evidently, something is propositional if and only if it consists of propositions. Propositions are, as we have seen,<sup>61</sup> objects of propositional attitudes. They can be true or false, and we can express them relative to contexts by ‘that’ clauses (Williamson 2000a, p. 194). Hence our first premise says that something is evidence only if it is a proposition. Intuitively, that claim might strike us as a bit odd, because we often talk about evidence with the intension of referring to things or sensations. For instance, evidence in a court trial may consist in specimens of DNA, while sensations of pain can be evidence for the belief that your leg is broken. Why should we believe that evidence must be propositional?

Williamson starts his argument with the “familiar claim” that only propositions can be reasons for belief, which is endorsed by some contemporary philosophers (Davidson 1983; Unger 2005, p. 204-6).<sup>62</sup> The claim is an offspring of Wilfrid Sellars’ attack on *the Myth of the Given* (Pryor 2005, p. 183; Sellars 1991, p. 164-70).<sup>63</sup> Sellars’ original idea was that beliefs are the only thing that could serve as justifiers. His argument can be summarized as:

- (1G) In order to be a justifier, you need to have propositional content represented assertively.
- (2G) Only beliefs represent propositions assertively.
- (3G) ∴ Only beliefs can be justifiers.<sup>64</sup>

Since it is plausible that we can enjoy experience with a propositional content *p* without believing *p*, the above argument is usually qualified by a *Premise Principle* (PP), which requires that justifiers “assertively represent propositions” in such a way that they “could be used as premises in an argument for that belief” (Pryor 2005, p. 189).

We can identify Williamson’s first premise with 1<sub>G</sub>, so our *PP*-version of Sellars’ argument cannot support that premise. Williamson’s strategy is therefore to single out theoretical functions central to the ordinary concept of ‘evidence’, and then ask what serves

---

<sup>60</sup> See Williamson (2000a, p. 193).

<sup>61</sup> See chapter 1.

<sup>62</sup> Actually, Unger’s claim is that reasons must be “propositionally specific”. Reasons need not be the same objects as the objects of knowledge and belief. His point is rather that there must be a certain parallel between reasons and the objects of knowledge (2002, p. 204-5).

<sup>63</sup> Davidson joined Sellars’ attack on the Given when he acknowledged that: “The relation between a sensation and a belief cannot be logical, since sensations are not beliefs or other propositional attitudes. What then is the relation? The answer is, I think, obvious: the relation is causal. (...) But a causal explanation of a belief *does not show how or why the belief is justified*” (1983 p. 143).

<sup>64</sup> See argument in Pryor (2005, p. 188). Pryor qualifies “beliefs” by saying “or other states that are epistemically like beliefs”.

them (Williamson 2000a, p. 194). One function evidence has is to be explained by hypotheses. According to Williamson, what hypotheses explain must be propositional, and hence we may conclude that evidence is propositional too (2000a, p. 195). Here's why: Hypotheses explain and provide answers to Why-questions. That hypotheses explain propositions is indicated by the fact that any expression of a sentence '\_\_\_because...' would be ungrammatical and ill-formed unless we fill the empty slots with declarative sentences (i.e. complements for 'that') (2000a, p. 195). We cannot simply explain *Albania* or *a bloodied knife* by saying something like, "Albania because ...", or, "Bloodied knife because ...". What a proper explanation seems to require is declarative sentences on both sides of the explanation-relation. For instance, one could say, "The bloodied knife connects Mr. Blue to the murder of Ms. Green because *that knife had Ms. Green's blood on it when it was found in Mr. Blue's house*", or, "I am moving to Albania because *I want to learn their language*".<sup>65</sup>

A similar point follows from considerations about probability: Often one compares the probability of two hypotheses  $h$  and  $h^*$  on  $e$  ( $P(h|h^*|e)$ ) by calculating the inverse probabilities of  $e$  on  $h$  and  $h^*$  ( $P(e|h/h^*)$ ) (Williamson 2000a, p. 195). The point is simply that in order to compare two competing hypotheses about evidence  $e$ , one often judges  $h$  to be more probable than  $h^*$  on the basis of what probability the evidence has on the hypotheses. What is important for us is the observation that nothing can be attached a probability unless it is propositional. After all, a probability is the probability *that ...* (2000a, p. 196). A third consideration that counts in favour of taking evidence as propositional is the fact that we use evidence to rule out hypotheses that are *inconsistent* with the evidence, and inconsistency is a relation between propositions (2000a, p. 196).

We choose our hypotheses on the basis of evidence, which is to say that we choose that hypothesis which is best explained, has the highest probability, and which is not ruled out on the evidence (Williamson 2000a, p. 197). These are three basic functions of evidence, and the above considerations indicate that all of them require propositions. Let us test the premise against some initial objections before we conclude anything. For one thing, it is often thought that perceptual experience is evidence of a non-propositional kind, but that seems to go against Williamson's first premise. In response, he grants that experience *provides* evidence without itself being propositional, but he adds that the evidence provided by experience *is* propositional (2000a, p. 197).<sup>66</sup> According to Williamson, what happens when experience license a belief, is still that some set of propositional evidence  $e_1, \dots, e_n$  is evidence for a

---

<sup>65</sup> I have italicised the hypothesis.

<sup>66</sup> As Williamson says, "to provide something is not to consist of it" (2000a, p. 197).

hypothesis  $h$ , but that  $e_1, \dots, e_n$  counts as evidence only because one is experiencing  $\epsilon$  (2000a, p. 197). In this sense, the evidential status of  $\epsilon$  is only mediated.

One reason for postulating this evidential indirectness of experience is that it can provide an elegant explanation of what goes wrong in unfavourable circumstances. In such cases, something only *appears* to be  $\gamma$  for S, and hence S fails to gain perceptual knowledge since the appearance is only illusionary. Importantly, S still undergoes some experience  $\epsilon$  in those circumstances, so if  $\epsilon$  has direct evidential status it is hard to see that we can avoid the *false* evidence that something is  $\gamma$ . On  $E = K$  that would constitute false knowledge, which goes against the factivity condition. But even without that condition it is somewhat awkward to talk about false evidence. If  $\epsilon$ 's evidential status is mediated we can explain such cases: Evidently one would still have perceptual evidence, but now  $\epsilon$  only confers evidential status on the true propositions that things appear to be thus-and-so (Williamson 2000a, p. 198).

It is still possible to come up with some intuitive counterexamples where a non-propositional state can count as evidence without being mediated. Take for instance a severe headache: It is not obvious that a headache is not non-mediated evidence for the belief that, "I have a headache", and so it seems that such a state could directly justify that belief (Pryor 2005, p. 192). For Williamson, that evidential-relation must be mediated, and he says that, "being directly aware *that* one is in pain is sufficient for both knowing and believing that one is in pain" (2000a, p. 59). Importantly, Williamson thinks that it is only by being related to the proposition that one is in pain that S can have evidence for the belief that one has a severe headache. The objector must be able to tell us how non-propositional states, like a severe headache, can be taken to justify, "I have a headache", rather than, "I don't have a headache". In order to choose between those beliefs, it seems that some *logical* relation is required between the evidence and the beliefs, and it seems plausible that only propositions can partake in such relations.<sup>67</sup> Hence none of these objections are strong enough to reject this premise.

Let us turn to Williamson's second premise  $E_{ii}$ : It says that a proposition  $e$  is included in S's evidence only if S knows  $e$ . This is a premise of pivotal importance for Williamson's position, since it makes knowledge a necessary condition for evidence and justification. Let us see whether we can provide it with sufficient support: Suppose that S reasons probabilistically from an observation base  $E$ , where all  $a_1, \dots, a_n$  were found to be  $\phi$ , to form the *true* belief that

---

<sup>67</sup> Pryor suggests that we can take states like headaches to have *certain logical properties* that rule out the latter belief and support the former, but it is unclear to me just what those logical properties could be if they are not something propositional (2005, p. 193).

$a_{n+1}$  will be  $\phi$  too. Intuitively, S did not know that  $a_{n+1}$  was  $\phi$  *prior* to the confirmation of her probabilistically justified belief. Now consider the following *false* hypotheses:

(h)  $a_1, \dots, a_n$  were  $\phi$ ;  $a_{n+1}$  was not  $\phi$ .

(h\*)  $a_1$  was not  $\phi$ ,  $a_2, \dots, a_{n+1}$  were  $\phi$ .<sup>68</sup>

According to Williamson, a natural thing to say is that  $h$  is consistent with S's evidence, and that  $h^*$  is not (2000a, p. 201). Notice that it follows from the fact that  $h$  is consistent with S's evidence, that S's true belief that  $a_{n+1}$  would be  $\phi$  is not included in S's evidence. Why is that so? Williamson's answer is that it fails to be a part of S's evidence because S does not know that  $a_{n+1}$  will be  $\phi$  (2000a, p. 201). That indicates that knowledge is a requirement on one's evidence. A related consideration is that *if* justified true beliefs (or some other good cognitive status short of knowledge) could be included in one's evidence, then a chain reaction may threaten to enlarge S's evidence to absurd proportions. For notice that *if* a set  $E$  justifies some further set of true propositions  $E^*$ , then  $E^*$  will also count as evidence.  $E^*$  will again justify a further set  $E^{**}$ , that will count as further evidence, and so on (2000a, p. 201).<sup>69</sup> That supports that knowledge is a *necessary* condition for evidence.<sup>70</sup>

It is now time to consider the third premise  $E_{iii}$ , which says that knowledge of  $e$  is also *sufficient* for  $e$  to be evidence. Some *prima facie* considerations favour that knowledge is sufficient: For one thing, there is a strong connection between warranted assertibility and knowledge.<sup>71</sup> It seems plausible that one's evidence is strong enough to warrant an assertion of  $e$  if and only if it is strong enough for one to know  $p$  (Williamson 2000a, p. 264). Since it is also plausible to take assertions as the exterior analogue of judgements, it follows that evidence is strong enough to justify one's judgement that  $p$  if and only if it is strong enough for one to know  $p$  (2000a, p. 238). A related consideration is that since belief seems to aim at knowledge, we can say that knowledge is the norm of belief (2000a, p. 47, 255-6). In this sense S's belief that  $e$  is only justified or warranted when S knows that  $p$ , and that indicates that knowledge is sufficient to justify one's beliefs.<sup>72</sup>

A worry about the third premise is that since much of one's knowledge is redundant, the internal evidential interconnections in one's body of evidence will be obliterated (2000a,

<sup>68</sup> See Williamson (2000a, p. 200). Also see chapter 4.

<sup>69</sup> Peter Unger provides linguistic evidence for the claim that evidence for S (or reasons) must be something that S knows, since it follows from  $e$  being evidence for S that: (i)  $e$  is true; (ii) S believes  $e$ ; and (iii) S is (absolutely) certain that  $e$  (2002, p. 206-14). Note that (iii) is not something Williamson takes to be an entailment of knowing.

<sup>70</sup> The second premise also implies that if  $e$  is included in S's evidence, then  $e$  is both something true as well as something that S believes. This is evident since S knowing  $e$  entails both these relations.

<sup>71</sup> See chapter 4.

<sup>72</sup> As Williamson puts it, "one should judge (or believe  $p$ ) only if one knows  $p$ " (2000a, p. 11).

p. 204). Recall *EV*: On that principle *S*'s evidence includes *e* when *e* is evidence for a hypothesis *h* and thereby raises the probability of *h*. But that requires that *e* is given a probability less than 1 on the relevant distribution of probability. Therefore, *EV* permits that a proposition *e* can be evidence for another in a non-trivial sense, even though *e* is in one's evidence. One can also prove that if  $E = K$ , then the evidential interconnections in one's body of evidence will be symmetric within one's body of knowledge due to *EV*. Williamson puts it this way: "[I]f *p* is evidence for *q* for one then *q* is evidence for *p* for one" (2000a, p. 204).<sup>73</sup> Rather than revising *EV*, we can invoke a notion of *independent evidence* and say that *e* is independent evidence for *h* for *S* only if *S*'s belief in *e* does not essentially depend on inference from *h*. That might be sufficient to put the worry of obliteration aside.

This ends the argument for  $E = K$ . The above considerations initially support that we take the three premises to be true, and thereby provide *prima facie* support for the soundness of Williamson's argument (2000a, p. 9). However, I will not make any conclusions now, since the subsequent section will critically examine the equation.

### 3.2 Objections and replies to the equation

In the following section I shall attend to some objections and worries about the equation  $E = K$ , and sketch initial replies to them. To begin with, one could worry that by equating *S*'s total body of evidence with *S*'s total body of knowledge, *S* is not always in a position to know what her own evidence is due to the anti-luminous character of our minds.<sup>74</sup> The extent to which this is a worry, depends on one's conception of rationality. Evidently, the rational agent is required to respect her evidence and conform her beliefs to it, but the question is whether that requirement entails that the agent must always be able to know what her evidence is (Williamson 2000a, p. 15). As a consequence of anti-luminosity, Williamson embraces a conception of rationality where one is not always in position to know what one is rationally required to do.<sup>75</sup> Hence this objection is not necessarily worrisome. Rather it might be endorsed since it fits well with the above conception of mind. That conception we have supported by independent considerations, so this objection might actually be taken to support the equation as it fits in with the overall picture (2000a, p. 16).

A related concern is that if  $E = K$ , then two persons may be internal duplicates and yet fail to have the same evidence since it is clear that two internal duplicates may differ in what

<sup>73</sup> I omit the proof, but see Williamson (2000a, p. 204).

<sup>74</sup> See the previous chapter. One's evidence will be anti-luminous because one's total body of knowledge are mental states.

<sup>75</sup> The conception of rationality that requires of *S* that she must always be in a position to know what her evidence is has sceptical implications. I discuss that in chapter 6 where I also provide an argument that goes against this conception of rationality.

they know (Harman 2002, p. 417).<sup>76</sup> Then what it is rational for  $S^{BAD}$  to believe in a bad situation may not be rational for her internal duplicate  $S^{GOOD}$  to believe in a good situation. According to Gilbert Harman, that feature of  $E = K$  leads to the worry that  $S^{BAD}$  is so radically impoverished that the chances of rational recovery are slim (2002, p. 423). After all, without a body of knowledge,  $S^{BAD}$  cannot be justified to believe anything at all since knowledge is necessary for justification. Closely related to Harman's objection is Williamson's observation that nothing he can say to the dedicated sceptic gives her reason to rationally come to believe what he says. That follows from  $E = K$  and the fact that the dedicated sceptic has already given up all knowledge, for in so doing she also gives up all evidence that could serve to justify such beliefs (2000a, p. 27). All this could be a worry to Williamson if refutation of scepticism was supposed to rationally recover the dedicated sceptic or  $S^{BAD}$ . But, arguably, such a task concedes too much to the sceptic. As Williamson puts it: "If a refutation of scepticism is supposed to reason one out of the hole, then scepticism is irrefutable" (2000a, p. 27). As a consequence, the most one can hope for is something which can prevent the potential sceptic from falling into the sceptical hole in the first place (2000a, p. 27). The price of Williamson's kind of evidence-model is therefore that it seems hopeless for these radically impoverished subjects to fully recover in a rational manner. That might be a cost, but then again this might just be one of those things one should relinquish in the first place.<sup>77</sup>

A third worry is the threat of evidential *regress*: It seems that evidence  $e$  of some  $h$  requires some other evidence  $e^*$  to support  $e$ , and that this  $e^*$  again requires some further evidence  $e^{**}$  in order to justify  $e^*$ , and so on *ad infinitum* (2000a, p. 191). In response, Williamson distinguishes between two senses of justification: Something can be *explicitly* evidence-based if it is influenced by prior beliefs about the evidence  $e$  for  $h$ , and it can be *implicitly* evidence-based if it is appropriately causally sensitive to the evidence  $e$  for  $h$ . Suppose a belief  $h$  is explicitly evidence-based because it is influenced by  $e^*$ , where  $e^*$  is evidence for  $e$ . This is possible even though  $e^*$  is not explicitly evidence-based on some further belief  $e^x$ . If the minimal requirement on evidence and justification is only this implicit evidence-base, then the threat of regress evaporates. It is plausible that some belief  $h$  can be causally sensitive on its evidence  $e$  without any need of further beliefs about the evidence for  $h$  (2000a, p. 191-2). Let me briefly illustrate: One's belief that, "I have a headache", is firmly evidence-based by the evidence that, "My head is aching". It seems very awkward to require further beliefs as explicit evidence-base for the latter proposition.

---

<sup>76</sup> We argued this in the previous chapter. Notice that this is consistent with the claim that knowledge and evidence *supervene* on  $S$ 's mental states, since mental states are taken to be wide and externalist (Williamson 2000a, p. 191).

<sup>77</sup> I will return to some related issues when we encounter the sceptic more directly in chapters 5 and 6.



This last point is related to Williamson's modest kind of foundationalism on which all one's knowledge serves as the foundation for all one's justified beliefs (Harman 2002, p. 422; Williamson 2000a, p. 186). Let me end this section by emphasizing that feature of his account: We have already seen that the regress of justification seems to end at knowledge, now this point can be cashed out as a distinction between *absolute* and *relative* justification. Often a belief is justified relatively to some other beliefs from which it has been derived. That was the case with  $h$  when  $h$  was explicitly evidence-based on some  $e^*$ . But we can also say that a belief has *absolute* justification if and only if it is justified relative to one's knowledge (2000a, p9). That was the case with the implicit evidence-base that  $h$  had on  $e$ . As we have seen, this makes justification possible only for those who have some knowledge, and that must be captured by our conception of the point of justification. The primary point of justification is not to serve as one of the necessary and sufficient conditions for knowledge, rather justification is primarily a status which knowledge can confer on other beliefs (Williamson 2000a, p. 9). Hence there is an important sense in which knowledge figures in Williamson's account primarily as what justifies, and not as what gets justified (2000a, p. 9).

### **3.3 Inexact knowledge and margins for error**

There are some important limits to what we can come to know that follow as consequences of a conception that takes knowledge to be a purely mental state in a mind that is characteristically anti-luminous. Ignorance is a human state that is just as natural as knowledge is, so we ought to take some precautions in order to secure that what we take ourselves to know is really known. One natural thing to do is to require of knowledge that it leaves a big enough *margin for error* to satisfy a reliability-constraint (Williamson 1996, p. 216-7). This section will briefly explore these precautions.

Suppose you walk into a hall and find it filled with a crowd of a large number of people. You gaze at the crowd, and approximate their number. Perhaps you know that there are less than a thousand people but more than 50 in the hall. Suppose the exact number of people in the hall at the relevant time is  $n$ , and suppose further that you actually believe that there are  $n$  people in the hall. Do you know that there are  $n$  people in the room? Arguably, you do not and cannot know this. Because notice that on your means of investigation (the rough gaze), there could easily be the case that the hall was filled with  $n+1$  people without you noticing the difference. Of course, one could count, but that is not the point. The point is, that most of the time one perceives *roughly* and not exactly, and that there are cases so similar to

the actual case that one wouldn't notice the difference (Williamson 1996, p. 217). What you know in this case and in similar situations is what we call *inexact* knowledge.

Inexact knowledge is an important phenomenon, since its existence imposes restrictions and limits on our knowledge and beliefs. Evidently, beliefs must be reliable in order to be known (Williamson 2000a, p. 100). In the above crowd-case we saw that one's belief that there are  $n$  people in the hall was not reliable, since there were very similar cases where the number of people was not- $n$  and in those cases one's belief would be false. For a belief to be knowledge, it is necessary for it to have a *margin for error*. Let us say that a belief that  $p$  has a margin for error in a case  $\alpha$  only if the condition  $p$  obtains in all similar cases  $\alpha_1, \dots, \alpha_n$  (1996, p. 226).<sup>78</sup> Hence a belief that  $p$  obtains can be knowledge only if  $p$  obtains in all sufficiently similar cases (1996, p. 227). This is captured in the following principle:

(Reliability) If S believes  $p$  truly in a case  $\alpha$ , one must avoid false belief in other cases sufficiently similar to  $\alpha$  in order to count as reliable enough to know  $p$  in  $\alpha$ .<sup>79</sup>

We can now see that anti-luminosity fits very neatly into this picture as a special case of inexact knowledge. Suppose that you are in a case where you want to judge whether you are in some mental condition  $\theta$ . Then what you consider is whether  $\theta$  obtains (in yourself, as it were). In order for your belief to be knowledge you must be in a case where the belief that  $\theta$  obtains has a margin of error. In other words,  $\theta$  cannot fail to obtain in any sufficiently similar case. Conversely, one will not be able to know that  $\theta$  obtains when there are sufficiently similar cases where  $\theta$  does not obtain. The anti-luminosity of mind is just the claim that none of our beliefs about our own paradigmatic mental states will have the required margin for error when the condition is near the boundary where it no longer obtains, and thus that those beliefs in general are not reliable enough to constitute knowledge. They are anti-luminous.

Another special case of inexact knowledge is when what one believes to obtain,  $p$ , is of the form, "I know that  $p$ ". Again, the belief that oneself knows that  $p$  ( $BK(p)$ ) requires a margin for error if it is to be known ( $KK(p)$ ). In a now familiar way that margin will require that the condition  $K(p)$  obtains in all cases sufficiently similar to the one where S  $BK(p)$ . Since knowledge is a mental state and hence anti-luminous, it follows that there are some cases  $\alpha_1, \dots, \alpha_n$  where S  $K(p)$  obtains and yet there are sufficiently similar cases where it does not obtain. Importantly, in those cases  $\alpha_1, \dots, \alpha_n$  S will know  $p$  and yet be unable to know that she

---

<sup>78</sup> The degree and kind of required similarity may depend on the circumstances. Generally, the more accurate the cognitive capacities are, the narrower is the margin needed to achieve a given level of reliability (Williamson 1996, p. 226).

<sup>79</sup> See Williamson (2000a, p. 100). We used a version of *Reliability* in our discussion about luminosity in chapter 2.

knows  $p$  (Williamson 1996, p. 228). That is a failure of the so-called  $KK$ -principle, which says that if  $S$  knows  $p$  then  $S$  knows that she knows  $p$  (Williamson 1996, p. 223).

Notice how knowledge of knowledge requires two margins of error: One for  $K(p)$ , and another for  $KK(p)$ . Thus there is an important sense in which there are two different circumstances that justify one's saying  $K(p)$  and  $KK(p)$ .<sup>80</sup> We can now generalize that point: Since every iteration of knowledge will require a new margin for error a new iteration will widen the margin for error. Hence by making such iterations one makes one's belief more reliable. Thus iteration is a natural way to secure that one's beliefs are true (Williamson 1996, p. 228). As Williamson puts it, "If truth is a hit, knowledge is a safe hit" (1996, p. 228). We can add that knowledge makes that hit safer by every new iteration.

### 3.4 Conclusion

In this chapter we have seen how Williamson uses the concept of knowledge to elucidate crucial features of the notions of evidence and justification, which are important notions in our conceptual scheme. Williamson equates one's body of evidence with one's body of knowledge, and uses that equation as his foundation for further justification of beliefs. Importantly, justification is something conferred by knowledge on other beliefs, and not something that is meant to justify knowledge in the first place. Hence the primary role for knowledge in our conceptual scheme is to justify. In this sense, the regress of justification stops at knowledge, and we should make sure that we resist any temptation to give that foundation up. Since there can be no justification without knowledge, the sceptic, who has already given up all knowledge in exchange for scepticism and ignorance, is in a very impoverished condition to which there is no immediate hope of rational recovery.

Another side of the coin is that the equation of evidence with knowledge coupled with the externalist conception of mind where knowledge is a mental state commits Williamson to an externalist conception of justification and evidence. If one's evidence is one's knowledge, and knowledge is a mental state that constitutively depends on the environment, then it cannot be the case that  $S$  at any time she holds a justified belief  $P$  is in a position to know that  $P$  is justified (Steup 2006). Because the justification of our judgements depends on the mental condition of knowledge, we are not always in a position to know that our judgements are justified. The exterior analogue to judgement is assertion, so a treatment of that issue suggests itself. The next chapter will explore how knowledge can also be used to elucidate that notion.

---

<sup>80</sup> Hintikka has a different intuition here and thinks that the required circumstances are the same, and so he defends a version of the  $KK$ -principle. He makes sure that he qualifies his endorsement, by saying that: "In the primary sense of *know*, if one knows one *ipso facto* knows that one knows" (1962, p. 23). On Williamson's conception, there is no need to distinguish between primary and secondary senses of know.

## Chapter 4: Assertion

*No sort of speech act is as important for philosophers to understand as assertion.*

Robert Brandom

*Only knowledge warrants assertion.*

Timothy Williamson

According to Frege, a full account of language must deal with three features of language: (i) Reference (*Bedeutung*); (ii) Sense (*Sinn*); and (iii) Force (*Behauptende Kraft*) (Dummett 1981, p. 295; 1993, p. 204).<sup>81</sup> Arguably, to know the sense of all expressions in a language (*L*), and thereby their reference,<sup>82</sup> is not enough to grasp the *significance* of uttering expressions in *L* or the *point* of doing so, and that requires the explanatory scope to include *force* (Dummett 1981, p. 295). A basic utility of language is that we can *use* it to *do* different things. One such thing is the making of *assertions*. According to Williamson, we can elucidate that practice by employing the notion of knowledge. Assertion is the exterior analogue of judgement, so an intuitive idea is to think of assertions as regulated by a norm of knowledge in a way that runs parallel to the justification norm on judgement (2000a, p. 238). The following chapter will explore those aspects of Williamson's account.

When we investigate assertion we query about the nature of assertoric force. We could ask "the more ambitious" question about the non-circular necessary and sufficient conditions for making assertions, but Williamson restricts himself to only ask about the constitutive rules (2000a, p. 239). We shall follow his lead, so in section 4.1 I commence our discussion by clarifying the notion of *constitutive rules*, while in section 4.2 I discuss the structure of a *simple* account. It is methodologically credible to seek first and foremost for an account based on one unique rule, and thus to search for an account that is simple. A candidate for such an account is the simple truth account, which is discussed and rejected in section 4.3. Williamson defends the idea, originally endorsed by Peter Unger, that only *knowledge* governs assertion.<sup>83</sup> In section 4.4 I articulate Williamson's *knowledge account*, whereas the following sections (4.5-4.8) discuss its plausibility. In section 4.5 I consider an argument from *probability*, while 4.6 and 4.7 discuss support from ordinary *conversational patterns* and *Moore's paradox*.

---

<sup>81</sup> The terms in parentheses are Frege's original terms (1892; 1918; 1949). Frege illuminates his terms accordingly: The reference of a proper name is the object itself that is designated thereby (1949, p. 201); while for a sentence the reference is to its truth-value (the True or the False) (1949, p. 203). The sense of an expression  $\phi$  is the manner in which the designation is presented (1949, p. 200).

<sup>82</sup> Frege: "there corresponds a definite sense to the sign and to this sense there corresponds again a definite nominatum" (1949, p. 200).

<sup>83</sup> Unger's version of a knowledge account of assertion differs slightly from Williamson's, but both agree about emphasizing the impropriety of asserting *p* without knowing it (Unger 2002, p. 250-71). See n115 below.

#### 4.1 Illocutionary acts and constitutive rules

Assertions are acts where an asserter (S) claims that something holds. As a consequence, assertions function naturally to express cognitive attitudes like beliefs and knowledge (Pagin 2007). Austin thought of assertions as *illocutionary* acts, which are the kind of speech acts that are done *in* uttering words. The successful achievement of those acts depends on the proper *attachment of force* to one's utterance (Austin 1963).<sup>84</sup> Importantly, speech acts are rule-governed in such a way that we can say that the rules constitute the acts. This also goes for assertion, so in this section we will discuss what it means to say that assertions are essentially rule-governed, and I will also examine the nature of those rules.

Typically, we assert something by uttering a *declarative sentence*, which says that so-and-so is the case (Pagin 2007). Grammatically, the verb 'assert' takes a 'that' clause as a complement, so when S asserts that  $p$ ,  $p$  is replaced by a declarative sentence. We have already seen that propositions are expressed relative to contexts by 'that' clauses, so it is a small step to say that assertions relate S to propositions (or propositional contents) (Williamson 2000a, p. 195). In this sense, assertions are *propositional acts* (Pagin 2007). It is tempting to say that *all* illocutionary acts are constituted by the combination of some force  $F$  with a proposition  $p$ , and that this unit  $F(p)$  is the *minimal unit* of human communication (Searle and Vanderveken 1985, p. 1). But I will not argue this point, since it is not important for what follows. Note that the types of illocutionary acts correspond to the type of illocutionary force. In order to instantiate a speech act, like for instance *swearing* or *promising*, we must attach *commissive* force ( $\perp$ ) to a proposition ( $p$ ). Then we can symbolize that S promises to do  $\phi$  at  $t$  as:  $\perp(S, p)$ . Here ' $\perp$ ' symbolizes the force of commitment that is attached to S, while ' $p$ ' abbreviates that S does  $\phi$  at  $t$  (Searle and Vanderveken 1985, p. 192).

An important insight of Austin's is that one brings about the proper attachment of force only if one makes a locutionary act in *appropriate circumstances* (1979, p. 131). If one utters expression  $\phi$  in a context C,  $\phi$  will *count as*  $\phi$ -ing if and only if a set of conditions  $\Delta$  is satisfied. Importantly, those speech acts can be subject to *infelicities*; it might not be appropriate to assert  $p$  in all possible circumstances. In an important sense, illocutionary acts can "fail to come off in special ways" and be "unhappy" (1979, p. 132). A natural thought is to think of these infelicities as the result of breaking certain transparently simple rules (Austin

---

<sup>84</sup> *Illocutionary* acts are at the mid-level between *locutionary* and *perlocutionary* acts (Austin 1963; Pagin 2007). A locutionary act is just the act of "saying something" with a particular *meaning* (Austin 1963, p. 94; Pagin 2007; Saul 2005); while a perlocutionary act is something we do by means of illocutions and which depend entirely on the hearer's reaction (Pagin 2007). We can say that they are the act of uttering words with a particular *effect* (Saul 2005). Illocutions are the acts done *in* uttering the words (Saul 2005). Austin's distinctions are not entirely sound, since the illocutionary acts depends on the hearer in the sense that they require the "securing of uptake" (Austin 1963, p. 116; Pagin 2007). But assertions obviously fail such a requirement since one can assert something without securing uptake.

1979, p. 132). As a consequence, we may say that speakers make themselves liable to praise or criticism by performing speech acts. Typically, speakers are evaluated depending on the level of conformity with the rules, but this is not unique for speech acts since we may use rules to regulate all sorts of behaviour. But, as we shall see, language-use is essentially rule-governed, and that, in turn, has ontological consequences. The rules of speech acts have a bearing on the very *nature* of the acts; they constitute them (Williamson 2000a, p. 238).

Consider the act of jumping: We may regulate that act by rules. This is already done to some extent, since jumpers incur resentment as well as admiration. However, the rules that pertain to jumping do *not* constitute that act. In contrast, there is an important sense in which the rules of language *constitute* speech acts and enable linguistic behaviour in the first place. Can we explain this ontological difference? According to Searle, we can characterize that difference in terms of a distinction between two types of rules: *Regulative* rules (R-rules) and *Constitutive* rules (C-rules) (1969, p. 33). Searle invokes these notions in order to characterize what he takes to be an irreducible difference between linguistic and other kinds of behaviour.<sup>85</sup> In contrast to acts like fishing or jumping, that have ends-means relations depending only on matters of physical fact, speech acts in general (and illocutionary acts in particular) are matters of convention. Consider a case where S does  $\phi$  in order to catch a fish: If S catches a fish by so doing it is very awkward to say that it was a matter of convention (Searle 1969, p. 37). In contrast, speech acts performed within a language *are* matters of convention: If S does  $\phi$  (e.g. producing the sound “I promise to go fishing tomorrow”), then that is a physical fact that *counts as* doing a particular speech act  $\gamma$  (e.g. I *promised* to go fishing) as a matter of convention (Searle 1969, p. 37). In the same manner, rules for assertion are said to *constitute* the act. Assertions are more like chess than fishing, since neither of them can *exist* without their rules.

The R-rules regulate forms of behaviour whose existence is logically *independent* of the existence of rules (Searle 1969, p. 33-4). Fishing, jumping, and producing sounds like, “I know your name”, provide good examples of acts that we could regulate with R-rules. R-rules are often basis for appraisals, and are typically formulated in phrases like, “Do *X*”, or, “If *Y* do *X*” (Searle 1969, p. 34, 36). In contrast, C-rules are regulations that *create* or *define* new forms of behaviour whose existence is logically *dependent* on the existence of rules. In order to define new forms of behaviour, systems of C-rules must include rules that are formulated as, “*X counts as Y*” (Searle 1969, p. 33-4). According to Williamson, rules are constitutive

---

<sup>85</sup> Searle takes C-rules to be ineliminable and irreducible since some illocutionary verbs are *not* definable in terms of intended perlocutionary effects (i.e. in terms of its illocutionary point), even though other illocutionary verbs are. Searle says that he sides with *institutional theories* of communication as opposed to reductionist and naturalistic theories of meaning (1969, p. 71).

only when they are “essential to that act” (2000a, p. 239). The C-rules are essential in the sense that they *necessarily* govern all performances of the act. This follows from the fact that if a set of C-rules ceased to govern some act  $\phi$ ,  $\phi$  would cease to exist too. Note that the necessity of C-rules is *not* that they are necessary conditions for actual performance of the act; rather their necessity is normative in the sense that they *must* govern those acts in order for them to exist, and hence that they govern them unconditionally (2000a, p. 239-40).<sup>86</sup>

Let me illustrate with an analogy: When S cheats in chess S does not cease to play. The same goes for speech acts: When S makes an assertion S is liable to criticism if she breaks the rule(s). However, one is *not* criticized for failing to assert, but for asserting when one is not supposed to (2000a, p. 240). Note that these normative features do not entail moral commitments. To break a rule in a game does not licence moral denigration unless *moral* principles are added. The same goes for C-rules. To say that the normativity of C-rules is moral is to put the cart before the horse; it is rather the case that moral blemish is possible due to the nature of these acts (Williamson 2000a, p. 240). Neither are C-rules clearly teleological, since it is possible to realize the main point of the act by breaking the rules, and, conversely, to miss the point by meeting the requirements. Consider another analogy: Suppose the point of soccer is to score most goals (internal *telos*) and to entertain one’s audience (external *telos*). It is possible to accomplish that *by* breaking the rules, as well as to comply with the rules without scoring any goals or entertaining the spectators (Williamson 2000a, p. 240).<sup>87</sup> Similarly, we can achieve both the internal and external *telos* of assertion by breaking the C-rule(s), as well as flout them by being in full compliance with the rules.

A final remark on the notion of C-rules is that any speech act *A* is constituted by a set of C-rules  $\Delta$  (one, some or many) in the sense that  $\Delta$  only governs the performance of *A*. Hence  $\Delta$  can serve to *individuate* the act. A different act *B* would necessarily have different rules, so *B* must be governed by a set  $\Phi$  that is not identical to  $\Delta$ . So when we find the C-rule(s) for assertion it is the *unique* act being governed by that set (Williamson 2000a, p. 241).

#### 4.2 The structure of a simple account

We have seen that some rules are constitutive of the speech acts they govern, and thus among the *necessary* conditions for the speech act. Williamson does not pursue the *more ambitious question* and ask for the necessary and sufficient conditions for assertion, rather he

---

<sup>86</sup> The difference is one of alethic versus deontic modality (i.e. reading necessity as “in accordance with a rule”). See Brandom (1994, p. 10).

<sup>87</sup> Consider Diego Maradona’s “Hand of God”. Maradona scored with his left fist in the World Championship semi-final between Argentina and England in 1986. Since the referee acknowledged the goal Maradona satisfied both internal and external *telos* by breaking the rules.

settles on the question of identifying the C-rule(s).<sup>88</sup> *Simplicity* is both a tempting option and a virtuous theoretical property, so one should first and foremost try to find a *simple account* of assertion. The general structure for such an account is that it must be based on *one unique rule* that constitutes the act (Williamson 2000a, p. 241). Consider the following:

(C-rule)                      One must: assert  $p$  only if  $p$  has C.<sup>89</sup>

The C-rule is in *imperative* mood and licenses the speaker to assert  $p$  only if  $p$  has property C, and so it unconditionally forbids the combination of S asserting  $p$  without  $p$  having the qualifying property.<sup>90</sup> To break the rule is of course possible, since it would be pointless to prohibit that which cannot be done. A simple account of assertion employs the C-rule as the *unique* and the *only* rule that governs assertion, which serves to *individuate* the act. It is necessary to master the C-rule in order to *grasp* assertion, which requires some sensitivity to the difference between compliance and disobedience to the rule (Williamson 2000a, p. 241).

Any account of assertion is committed to explain other considerations of normative character concerning the act. Simple accounts must derive these results from the joint outcome of the C-rule and considerations not specific to assertion in order to remain *simple* (Williamson 2000a, p. 241). Consider the following illustration: In certain contexts it has severe moral and juridical consequences to assert  $p$  when  $p$  lacks C. Suppose the correct C-rule is a truth-rule that strictly forbids false assertions. There are contexts where one can be held morally responsible for asserting  $p$  without  $p$  being true, which is insufficiently explained by the truth-rule alone. An option is to derive such moral sentiments from other principles not specific to assertion. Suppose the context is that one is a witness in a U.S. court trial, and one is sworn in by oath to tell nothing but the truth. To swear is a speech act of its own (a commissive) with its own set of constitutive rule(s), which commits the speaker to take a certain course of action (Searle 1979, p. 155). In our context that course of action is to tell the truth, so if one were to assert  $p$  without assertoric license it would also break one's oath. Hence the moral sentiments can be derived as the joint outcome of the truth-rule and one's commitment to tell nothing but the truth. This shows that a simple account based on the truth-rule could retain its simplicity in this case. An analogous derivation is required in cases where compliance with the C-rule meets moral condemnation. Consider an analogy: It is considered rude to keep on playing an obviously lost position in chess. That is not something we have

---

<sup>88</sup> An attempt to answer "the more ambitious question" is provided by Searle in his taxonomy of illocutionary acts (1979, p. 1-29).

<sup>89</sup> See Williamson (2000a, p. 241). The rule can be formalized in standard deontic terms as:  $P ( \vdash(p) \supset C(p) )$ . Read ' $P$ ' as a permission operator and our formulation says that assertion of  $p$  is permissible only if  $p$  satisfies C (Kahane and Tidman 2003, p. 431).

<sup>90</sup> To break the rule will constitute an *unhappy* performance (Austin 1979, p. 132).



regulated by chess-rules, rather it follows from social considerations.<sup>91</sup> Importantly, this goes for assertion too: Perhaps the assertion of, “Pierre has put on weight lately”, is true and hence warranted on the truth-rule, yet one risks being impudent by asserting it in certain contexts.

We have seen how speakers receive criticism for performing appropriate or inappropriate speech acts. Importantly, propriety is a matter of salience. Let me illustrate: Sometimes it is appropriate according to the C-rule to assert  $p$ , and yet the assertion is inappropriate in the context. We can make a distinction between salient and non-salient senses of appropriate assertion (Williamson 2000a, p. 241). When it is saliently appropriate to assert  $p$ ,  $p$  satisfies the C-rule. Conversely, when it is non-saliently appropriate to assert  $p$ ,  $p$  fails to satisfy the C-rule and yet the assertion is right. Notice that since a simple account of assertion should be able to tell why assertions sometimes go awry in a non-salient sense through derivations, we have a strong signal that our account is insufficient when this cannot be done.

### 4.3 The truth account

The bottom line for a maximally simple truth account (T-account) is that assertion is best served by the following truth rule (T-rule) that strictly forbids us to make false assertions:

(T-rule)                      One must: assert  $p$  only if  $p$  is true.<sup>92</sup>

The simple T-account is committed to explain all normative considerations as derivative of the T-rule or as the joint outcome of the T-rule and concerns not specific to assertion (Williamson 2000a, p. 242). There are some important objections to this account: For one thing, assertion is *not* the unique speech act to be governed by truth. To *conjecture*, *swear*, *predict* and *inform* that  $p$  all aim at truth, and hence it seems that truth-governance cannot serve to individuate assertions (Williamson 2000a, p. 244). Since a T-account is required to discern the unique set of C-rules that governs assertion, we need to show that assertion is more intimately related to the T-rule than any other speech act. The problem is that it is difficult to find support for this claim.

Perhaps we could say that assertions are more intimately connected to the T-rule, because we require more stringent evidence for assertions than of other speech acts? That suggestion works against the act of *conjecturing*, since to conjecture  $p$  can be acceptable on some evidence  $E$  that does not warrant assertion. However, the act of *swearing* might count against this suggestion since to swear  $p$  seems to require more evidence than assertion

---

<sup>91</sup> I owe this example to Torfinn T. Huvenes.

<sup>92</sup> See Williamson (2000a, p. 242). We can formalize the T-rule as:  $P (\vdash p) \supset \text{True}[p]$ .

(Williamson 2000a, p. 244). By parity of reasoning we could show that swearing is more intimately related to truth, and thus that the T-rule is *not* constitutive of assertion. As a response, one could attach individual evidential norms to all speech acts that aim at truth and thereby differentiate them, but this will strain the simplicity of the account (2000a, p. 244). What this suggests is that some evidential account may serve us better since it would have much less difficulty individuating the acts (Williamson 2000a, p. 245). Perhaps assertion is related to these speech acts as the genus is related to its species? On Searle's taxonomy, assertion is the *primitive assertive* and all other assertives are explained as variants (1985). A counter-suggestion could therefore be that other speech acts may be regulated by the T-rule due to this relation. This is possible so let us consider a stronger objection.

There is a plausible link between epistemic norms and assertion (Soames 1999, p. 31).<sup>93</sup> On the T-account such norms are secondary and provide evidential basis for one's satisfaction of the primary truth-norm, so the T-account must explain these norms derivatively on the T-rule.<sup>94</sup> Assume for the sake of *reductio ad absurdum* that the T-account is correct. Then it should be possible to derive evidential norms from the account (Williamson 2000a, p. 245). On the T-account, the T-rule is the unique rule governing assertion, so these *E*-norms must be derivative. One should be able to derive something like: "One *ought* to have evidence for one's assertions *because* they ought to be true" (2000a, p. 245). The desired derivation is captured in the following valid principle that holds for any act of  $\phi$ -ing governed by truth:

(Evidence (Ev)) If one must ( $\phi$  only if  $p$  is true), then one should ( $\phi$  only if one has evidence that  $p$  is true).<sup>95</sup>

*Valid conditionals* (like *EV*) are corresponding conditionals to logically valid arguments, so since *EV* is valid we are warranted to infer the consequent of it from the satisfaction of its antecedent. In our case the antecedent is that, "one must ( $\phi$  only if  $p$  is true)", which is the assumed T-rule. Hence from our assumption we can derive the consequent that, "one should ( $\phi$  only if one has evidence that  $p$  is true)". To have evidence that  $p$  is true "just is evidence for that assertion", so the T-account implies that one should "not make an assertion for which one lacks evidence" (Williamson 2000a, p. 245). Notice that it is essential for our *reductio* that evidence to  $\phi p$  must vary with the degree of badness we attach to asserting  $p$  when  $p$  is false. Intuitively, to assert  $p$  when someone's life depends on the truth of one's assertion

---

<sup>93</sup> As Soames says, "The act of asserting that it is *true* that  $P$  commits one both to having good evidence for the proposition that  $P$ , and to the proposition that it is true that  $P$ " (1999, p. 31).

<sup>94</sup> We can say that the epistemic norms are among the "preparatory conditions", whereas the T-rule is an "essential condition" (Searle 1969, p. 64). Note that Searle also takes it that a sincerity condition of belief is among the necessary conditions for assertion.

<sup>95</sup> See Williamson (2000a, p. 245). The parentheses indicate the scope.

raises the burden of evidence. We would not require strict evidence when mere trivialities are at stake.<sup>96</sup> Let me illustrate: Suppose A accuses B of murder by asserting *falsely* that, “B killed C with an axe”. That is more important than if A accused B of impoliteness by asserting falsely that, “B gave me a certain look”. Naturally we require better evidence to secure the truth of murder-accusations than the latter kind of assertion. Thus:

(Evidence II) One should (proportion one’s evidence that  $p$  is true to the seriousness of  $\phi$ -ing  $p$  falsely).

If  $EvII$  is accepted as a charitable reading of  $Ev$ , the stage is set for a *reductio*. It is sufficient to provide a case where the epistemic burden weighs heavily on the asserter without a proportionate importance to explain the requirement. Consider the following lottery case: Suppose B has bought a lottery ticket in a large lottery where the draw has been held, but the winner is yet to be announced. It is obvious that B’s chances of winning are slim, so suppose you assert (flat-out), “Your lottery ticket is a loser”, relying only on strong probabilistic evidence without inside information. Now assume that B’s ticket *is* in fact a loser, so that what you asserted was something true. Despite the fact that you asserted something true there is a strong intuitive pull that B can feel offence with good reason when she discovers your probabilistic evidence-base. B could say, “You didn’t really know that”, and thereby criticize your speech act. Thus *high probability* is not sufficient for flat-out assertion.<sup>97</sup>

Observe that your false assertion is not an important matter for B since nobody risks harm by you asserting it. Of course, we could stipulate extreme scenarios so that one risks harming B due to a pathological obsession with lotteries, but to prevent that we can just counter-stipulate B’s mental and physical health. Importantly, our intuitions remain intact even though it is a matter of trifling importance *and* good health is fixed. We may conclude that the condemnation does not rest on epistemic burden proportionally related to the gravity of flouting the T-rule. Can it be explained by the T-rule itself? Recall that the T-rule only forbids false assertions, and thus cannot explain our case since that assertion is true. As a consequence, we reach an absurdity that supports the retraction of our assumption. We may therefore conclude that the T-account is false.

Perhaps we can save the T-account by an appeal to Paul Grice’s notion of *conversational implicature*? The suggestion is that B’s resentment stems from the frustrated entitlement to assume that you had inside information, since that is an implicature of your

---

<sup>96</sup> According to Williamson, a “charitable reading” of  $Ev$  is that, “the required weight of evidence for  $p$  will vary with the badness of  $\phi$ -ing when  $p$  is false” (2000a, p. 245-6). We capture such a charitable reading in  $EvII$ .

<sup>97</sup> The lottery case is from Williamson (2000a, p. 246). He adds that: “In any case, for whatever reasons, probabilistic bases are ordinarily taken to be inadequate for assertion” (2000a, p. 249). This generalizes the claim in an important manner.

assertion. B's criticism is therefore that you misled her to infer an off-beam implicature. A natural option is to say that your assertion flouted Grice's *maxim of Quantity*, which urge you to not "make your contribution more informative than is required" (Grice 1989, p. 26). Since it is clear to all conversational participants in the lottery case that the chances of winning are slim, the maxim urges you to not assert that, "Your ticket is a loser". Consequently, your assertion entitles B to work out a conversational implicature, for instance that one's evidence-base is not just probabilistic. The problem with this reply is that it would warrant the same procedure on an assertion like, "Your ticket is almost certainly a loser". But intuitively that assertion is all right, so a Gricean story predicts counterintuitive results (Williamson 2000a, p. 247). Another worry is that implicatures essentially allow *cancelability*<sup>98</sup>, but it is not reasonable that one could assert things like, "The ticket is a loser, and by that I don't mean to imply that I have inside information" (Grice 1989, p. 39; Williamson 2000a, p. 248).

A third line of defence is an appeal to parity reasoning: By asserting, "Your ticket is a loser", one does something for which one lacks evidential basis, since for each ticket in the lottery the probability is the same. Parity reasoning would warrant the conjunction of all tickets being losers, and since that assertion is false the resentment is explained on the T-rule. Williamson takes this proposal to be somewhat problematic, since the necessary resources for making the application is not brought to us by the T-account. It is possible that for each ticket there is adequate evidence for asserting, "That ticket is a loser", without there being adequate evidence for asserting their conjunction.<sup>99</sup> I think Williamson's reply underestimates the possibility of an independently supported principle that may license the required conjunction. Yet, perhaps what is right about it is that any such principle is untenable *unless* we equate evidence with *knowledge*.<sup>100</sup> Consider a case where one's evidence for each  $p$  has a probability of less than 1: Then a long chain of conjuncts  $p_1, \dots, p_n$  may all have individual probability above a certain threshold of evidence below 1, but collectively they may enjoy a probability below the threshold. Perhaps a conjunctive principle may hold *only* if the threshold of evidence is 1? If so, there is a small step to equate evidence with knowledge, and that makes it very tempting to replace the truth-requirement with knowledge.

---

<sup>98</sup> One can cancel an implicature either explicitly by saying so or contextually by uttering something in a context that makes it clear that the speaker is "opting out". Cancelability is an essential requirement on conversational implicatures. See Grice (1989, p. 39, 44).

<sup>99</sup> I.e. we may have warrant to say for all  $x$  that, "x is a loser", without having to say, "all x are losers".

<sup>100</sup> If we can support a Conjunction Principle (CP) for evidence and rational acceptability we can derive the intuitive result of the lottery-case as a joint outcome of the T-rule and CP. To prove CP we would need to prove something like: If each of the propositions  $\phi$  and  $\psi$  is rationally acceptable, so is  $\phi \ \& \ \psi$  (Douven and Williamson 2006, p. 755). At any rate, Williamson argues that CP is only tenable when the required acceptability is a probability of 1. Our above problem is really a version of Kyburg's Lottery Paradox: On a Sufficiency Thesis (ST) a proposition  $\phi$  is rationally acceptable if  $\Pr(\phi) > t$  (2006, p. 755). Combine ST with CP and one can derive a lottery paradox that shows that only propositions having probability 1 are rationally acceptable (2006, p. 760).

Our discussion warrants a strong suspicion towards the T-account: Assertion appears to be governed by something other than truth, and the lottery-case suggests a non-derivational evidential rule. Perhaps knowledge can do a better job? We will discuss that option next.

#### 4.4 The knowledge account

The T-account fell into discredit because of its inability to accommodate evidential norms associated with assertions. This is crucial since a maximally *simple* account must explain such norms derivationally or non-derivationally from its C-rule. Perhaps it is easier to defend that assertion is governed by an evidential rule? A candidate account of assertion is one that is based on a *knowledge rule* (K-rule). Since we are looking for a *simple* account, the K-rule must completely govern our use of assertions, and so it will be required to do all the explanatory work needed of a unique rule. Let us first articulate the view, and then see if it can fulfil our requirements in the subsequent sections.

The K-rule is a variant of a *warrant rule* (W-rule), which governs the propriety or permissiveness for making assertions. The notion of *warrant* will be used somewhat technically as something that is sensitive to the epistemic circumstances of an asserter (S), and which can license an assertion. By satisfying a W-rule S is warranted to make an assertion, and it will be *saliently correct* for S to perform the act (Williamson 2000a, p. 241). Let us use  $warrant_{ass}$  to denote this “schematic sense” of the word. S is warranted<sub>ass</sub> to assert only if she can satisfy the following W-rule, which can be turned into a K-rule by replacing:

- |          |   |
|----------|---|
| (W-rule) | One must: assert $p$ only if one has <i>warrant</i> to assert $p$ . |
| (K-rule) | One must: assert $p$ only if one <i>knows</i> $p$ . <sup>101</sup>  |

The K-rule  $warrants_{ass}$  S to assert  $p$  only on the condition that S knows  $p$ , and as such it is a W-rule (Williamson 2000a, p. 243). A *simple knowledge account* (K-account) takes the K-rule to be the *unique* rule that governs the speech act of assertion and serves to *individuate* the act (2000a, p. 241). The K-rule strictly *forbids* unwarranted<sub>ass</sub> assertions, and if the account is correct ordinary speakers are implicitly sensitive to the rule since that is required to master the speech act (2000a, p. 241, 243). The account explains truth-norms as derivative of the K-rule and the principle of factivity. The rule is required to explain other normative features of assertion as the joint outcome of the K-rule and considerations not specific to assertion (2000a, p. 242-3). The following principles capture the basic tenets of this account:

---

<sup>101</sup> See Williamson (2000a, p. 242-3). The K-rule can be formalized as saying:  $P(\vdash(p) \supset \text{Know}(p))$ .

- (Cor.) If S asserts that  $p$  and does know that  $p$ , that assertion is (semantically) *correct*.  
 (Incor.) If S asserts that  $p$  and does not know that  $p$ , that assertion is (semantically) *incorrect*.<sup>102</sup>

Peter Pagin thinks that Williamson employs warrant in a specifically *semantic\** way, so that by violating the K-rule one commits a *semantic\* error* even though one can be permitted to assert in non-semantic regards.<sup>103</sup> The difference between warrant and  $\text{warrant}_{\text{ass}}$  captures this distinction. The following captures Williamson’s uniqueness-claim:

(Unique ( $Un$ )) Utterance  $u$  by S is an assertion that  $p$  iff it is necessary that ( $u$  is correct iff S knows that  $p$ ).<sup>104</sup>

The K-account requires that a speaker (S) needs to be aware of the rule in order to be capable of making assertions. Williamson recognizes this point, but he abstains from elucidating this notion of sensitivity any further (2000a, p. 240). Pagin has objected that children or communities of speakers “who do not have any idea of rules”, or at least, “not of rules applying to assertions”, can make claims and statements without satisfying this necessary requirement (Pagin forthcoming). If true that would go against the K-account, but, as Searle has suggested, we must sometimes assume that a speaker acts in accordance with a rule, even though she may not state this rule nor be conscious of the fact that this is what she is doing (1969, p. 42). Hence the speaker’s know-how may only be sufficiently explained on the assumption that she knows (has acquired, internalised, learned) a rule (Searle 1969, p. 42).

#### 4.5 An argument from probability

We have encountered a lottery case where the fundamental point was that we lack evidential authority to assert a lottery proposition despite the fact that its truth is very highly probable on one’s evidence (Williamson 2000a, p. 249). We can draw a general conclusion if we alter that case so that it covers any proposition about any subject matter and not only propositions about lotteries. Such a case will provide the basis for the following claim: One has never  $\text{warrant}_{\text{ass}}$  to assert some  $p$  when  $p$  has less than probability of 1 on one’s evidence (2000a, p. 249). If we can justify that claim we get strong support for the K-account, since it is plausible that when  $p$  has a probability of at least 1 on one’s evidence, one knows  $p$ .

<sup>102</sup> See Pagin (forthcoming).

<sup>103</sup> The sense in which the rules that govern assertions are *semantic* proprieties (in contrast to pragmatic) is controversial. Where one draws the line between semantics and pragmatics will have important theoretical consequences, but since this debate is very complex I will not engage it. Let me just briefly say that the way I talk about semantic proprieties in this chapter might be in tension with “the traditional view”, where semantics is concerned with *what is said* and pragmatics with *what is implicated*. This is the “Gricean” conception, although whether Grice himself held such a view is questionable (Szabó 2005, p. 3). Importantly, it seems that conversational implicatures is within the scope of semantics. How about *force*? It is not obvious how one could locate *force* in this “Gricean” picture, and thus it is not obvious on which side of the semantic-pragmatic boundary we ought to locate the rules of assertion. Since I am not providing arguments to decide this question, I will just mark my use of the term ‘semantic’ with a star (\*) to indicate that for my part this matter is not settled.

<sup>104</sup> See Pagin (forthcoming).

Let  $p$  be a proposition whose truth-value an expert knows, but of which you know nothing. Suppose that the expert holds a lottery, where out of a million tickets you get one. Instead of announcing the winner, the expert hands out a slip of paper to each participant of the lottery. According to the expert, there is only one winner in this lottery, and only he/she is correctly informed in the sense that he/she has the true member of the pair  $\{p, \sim p\}$  scribbled down on his/her piece of paper. It is stipulated that these conditions are perfectly clear to all of the participants, and that none can share knowledge about their own slip of paper to other participants. If one receives a piece of paper saying, “ $\sim p$ ”, there is a million-to-one chance that  $\sim p$  is false, and hence that  $p$  is true. Despite massive probability for  $p$  being true one is intuitively not warranted<sub>ass</sub> to assert it (2000a, p. 250). These intuitions can be given additional support from the original lottery case by the following plausible biconditional:

$$(1) \quad \vdash (p \text{ iff my ticket is a loser}).^{105}$$

If one is in a position to assert  $p$  one would also be in a position to assert, “My ticket is a loser”. Let ‘a’ stand for  $p$  and ‘b’ stand for “my ticket is a loser”. It would be awkward to deny that warrant<sub>ass</sub> of both  $\vdash (a \equiv b)$  and  $\vdash (a)$ , should not warrant<sub>ass</sub>  $\vdash (b)$ . In the lottery case we have already proven that the last proposition cannot be asserted. Consequently, we cannot have warrant<sub>ass</sub> to assert  $p$  without flouting the biconditional. The moral is: Since we have no evidence independent of the very high probability that one’s ticket is a loser, we are not warranted<sub>ass</sub> in asserting  $p$  (2000a, p. 250). Observe that all we required of  $p$  was that it possessed a truth-value, and hence  $p$  serves as a mere *placeholder* replaceable by *any* proposition. We can therefore generalize the result: The argument strongly indicates that very high probability on one’s evidence (short of 1) is not enough to warrant<sub>ass</sub> assertions for (almost)<sup>106</sup> any kind of proposition (Williamson 2000a, p. 250).

The claim is further supported by considerations on transitivity-relations for assertoric warrant<sub>ass</sub>: It is very plausible that *if* some  $p$  has less probability than  $q$ , and one is warranted<sub>ass</sub> to assert  $p$ , then one is also warranted<sub>ass</sub> to assert  $q$  (Williamson 2000a, p. 251). Contrapositively, we can say that *if* one is *not* warranted to assert  $q$ , then it is *not* the case that there is some  $p$  with a probability less than  $q$  and one is warranted to assert  $p$ . Read ‘P(x|y)’ as the *probability of x on y*, and let  $e$  stand for one’s evidence, and these principles are:

<sup>105</sup> ‘ $\vdash$ ’ symbolizes that  $p$  has an assertoric force attached to it. In *Begriffsschrift* Frege adds a vertical judgement stroke to the horizontal content stroke; without the first stroke, what follows the marker is only “a mere complex of ideas”, but with that stroke it is a *judgement* to the effect that the speaker recognize  $p$  as true (1879, p. 52). An assertion is the exterior analogue of such a judgement.

<sup>106</sup> The argument does not rule out propositions for which one is “bound to have independent non-probabilistic evidence”. Cartesian propositions, like “I think, therefore I am” and “I exist”, are examples of that. But they are irrelevant for us (Williamson 2000a, p. 250).

(Transitivity (Trans))  $\forall p \forall q \{[(P(p|e) < P(q|e)) \& \vdash^{\text{warrant}}(p)] \supset \vdash^{\text{warrant}}(q)\}$ <sup>107</sup>

(Contrapositive (Contra))  $\forall p \forall q \{\sim \vdash^{\text{warrant}}(q) \supset \sim[(P(p|e) < P(q|e)) \& \vdash^{\text{warrant}}(p)]\}$

If we endorse these principles, our general conclusion can be cashed out accordingly: For any  $p$  with determinate high probability short of 1, we can always envisage  $q$  with a higher probability than  $p$  but still short of 1, and where intuitively  $q$  cannot be asserted due to a lack of warrant<sub>ass</sub>. In our lottery case  $q$  was not warranted<sub>ass</sub> on a million-to-one probability, which satisfies the antecedent of *Contra*. Thus any  $p$  with a lower probability than  $q$  is not warranted<sub>ass</sub>. This follows since the first conjunct of the consequent *is true*, thereby forcing *falsity* on the second conjunct on pain of rejecting this principle. Our lottery case indicates that nobody is warranted<sub>ass</sub> to assert *any* proposition with a probability short of 1.<sup>108</sup>

What kind of probability are we talking about here? An intuitive and attractive proposal is to require probability 1 *on one's evidence*. Recall that we equated evidence with knowledge in chapter 3, so this means that  $p$  is only warranted<sub>ass</sub> just in case it has probability of 1 on what one *knows*. Given *that* equation the lottery case provides support for the K-account, since the requirement on warrant<sub>ass</sub> *is* trivially satisfied by knowledge; after all what one knows trivially has probability 1 on one's evidence.<sup>109</sup> Since the only way to warrant<sub>ass</sub> assertion  $p$  is for  $p$  to have the probability of 1 on one's evidence, then, on the assumption of  $E = K$ , we can say that  $p$  is warranted<sub>ass</sub> whenever  $p$  is something one knows (2000a, p. 251).

Peter Pagin has questioned whether Williamson provides the correct diagnosis of the observation that propositions with a probability short of 1 are not warranted<sub>ass</sub>. According to Pagin, the lottery case exploits three important distinctions: (i) conclusive vs. inconclusive (probabilistic) evidence; (ii) asserting what is mutually known vs. asserting what is not mutually known; and (iii) assertions based on causal information vs. assertions based on indirect grounds (forthcoming). In the lottery case we entertain a proposition whose evidence is inconclusive and the assertion of which is mutually known *and* based on indirect grounds. The important point is that each of these vectors may be responsible for blocking warrant<sub>ass</sub>. In order to blame probability we need cases where these distinctions come apart (Pagin forthcoming). Let us start with (ii): Suppose you assert, "Your ticket is a loser" ( $\psi$ ), and that you know B is ignorant of the fact that more tickets have been sold than the organizers have announced. According to Pagin, one can now assert  $\psi$  on probabilistic basis *if* one thinks that

<sup>107</sup> Read ' $\vdash^{\text{warrant}}(p)$ ' as "warrant<sub>ass</sub> to assert  $p$ ".

<sup>108</sup> Observe that this is on a par with the result that no proposition with probability less than 1 are rationally acceptable (Douven and Williamson 2006, p. 760). We touched upon these issues in section 4.3 above.

<sup>109</sup> This follows on the EV-principle in the previous chapter: Since one's evidence is what one knows, then to know  $p$  will have probability of 1 since  $p$  will also occur among one's evidence. Any probability  $n$  on oneself has the value 1 ( $P(n|n) = 1$ ).



the real probability is so low that all hope is gone (forthcoming). I do not share his intuition: For one thing, there is no need to lose *all* hope since there is still a chance that the ticket wins, so if one thinks that *all* chance is gone when asserting  $\psi$  it is unwarranted<sub>ass</sub>.

Pagin also thinks that future tense versions of  $\psi$ <sup>110</sup> made before the draw are warranted<sub>ass</sub>, even though they are probabilistically based (forthcoming). According to Pagin,  $\psi$  flouts warrant because it concerns a past event about which one *ought* to have causal information in order to assert  $\psi$ . Future tensed assertions cannot in any case have causal relations to what they assert, so one is warranted<sub>ass</sub> to assert them without causal relations. But, as Hawthorne has argued, future tense assertions of lottery propositions are just as unwarranted<sub>ass</sub> as present tensed versions (2004, p. 21-31). Pagin is aware of Hawthorne's intuitions, but he thinks that the following conversation supports his own view:

- B: I'm not taking the job since I've bought a ticket in a lottery where the winner gets a million.  
A: Come on, your ticket will not win.

My own intuitions go with Hawthorne on this. I agree with Pagin that *something* goes wrong in B's reasoning, but it seems that A overacts and would be better off by saying something like, "Come on, you *don't know* that". What *is* wrong with B's reasoning is that B flouts the "intuitive link" between knowledge and practical reasoning (Hawthorne 2004, p. 29). B does not know the crucial premise in this inference, and that is why her reasoning is wrong.<sup>111</sup> So I think both objections from Pagin are flawed. We may conclude that the intuitions aroused by considerations of probability seem to be reliable, and that they make a good case for the connection between knowledge and assertion.

#### 4.6 Conversational patterns

Another line of support for the K-account is evidence from *conversational patterns*. Grice has taught us that the connection between meaning and use is indirect, so we must be somewhat cautious when we draw semantic implications from patterns of linguistic use (MacFarlane 2007, p. 204). Yet, conversational patterns are important evidence since they may reveal our "reactions and thoughts in ordinary conversational situations" (Unger 2002, p. 260). The following section will explore some of the evidence we have about assertions.

---

<sup>110</sup> E.g. "Your ticket *will not* win".

<sup>111</sup> Hawthorne cashes out the intuitive link accordingly: "That one does not know a lottery proposition prohibits one (...) from asserting it. It seems to prohibit one from using it as a premise in one's deliberations about how to act" (2004, p. 29).

Suppose S (incorrectly) asserts that, “Barry Smith is an Englishman” ( $\theta$ ), and it is obvious that the assertion is sincere.<sup>112</sup> It seems that it would be appropriate to respond by saying, “How do you know that?” The form of this question weakly *presupposes* that S has an affirmative answer, and, accordingly, the hearer *can* convey to S that she trusts S to actually know  $\theta$ . Notice that when one puts a sarcastic stress on “that” the question no longer presupposes an affirmative response. Could this question be justified unless knowledge warrants<sub>ass</sub> assertion? Evidently, one could not make an *implicit* challenge to S unless knowledge governs assertion (Williamson 2000a, p. 252-3).<sup>113</sup> The contrast is with questions like, “Where did you read that?” or, “When did Barry tell you this?” These questions *might* be an appropriate challenge to the effect that one asserts  $p$ , but they are *not implicit* since one can have warrant<sub>ass</sub> without being able to read and without having the chance to talk with Barry Smith. But our first question seems appropriate for all kinds of assertions.

A clearer challenge to S would be to ask, “Do you know that?” Note that the form here is possibly less polite and more aggressive than the first question, even though stress may again alter the tone. For suppose that one puts stress on “you” in a slightly surprised tone. Then it comes out as surprise rather than aggressive. Without this stress one *can* convey that one does not presuppose an affirmative response. Hence the question is open. On the K-account the potential aggressiveness is explained as the fact that only knowledge warrants<sub>ass</sub> assertion. The question can explicitly challenge S’s warrant<sub>ass</sub> to do what she has already done. If knowledge is *not* the only warrant<sub>ass</sub> S could in fact blamelessly respond with a simple, “No” (Williamson 2000a, p. 253). But, as Unger points out, S could never get off the hook by saying, “I never said I knew it” (2002, p. 264).<sup>114</sup> Again, that supports the K-account.

Consider another case from Unger (2002, p 260-1): Suppose that you hear a colleague (C) assert that you will get a substantial rise in salary. As it turns out, C did not know that you were going to get a salary although she was *fairly confident* of it. Evidently, C is not warranted\*<sup>115</sup> in making the assertion in question. Unger thinks that C’s degree of violation is “of a piece with lying” in so far as the act was intentional (2002, p. 261). *If* assertion is

---

<sup>112</sup> Barry Smith *is* a Scot so I hope this won’t insult him.

<sup>113</sup> It is an implicit challenge, since the questioner “politely grants that the asserter does know  $p$ ” by asking in a form that presuppose an affirmative answer. The questioner can suspect a non-affirmative answer (as one would with  $\theta$  since Barry Smith *is* a Scot), but a polite tone is secured by posing the question in a *presupposive* form (Williamson 2000a, p. 253).

<sup>114</sup> Unger use similar utterances and conversational data to support that one *represents oneself* as knowing  $p$  by asserting it. Unger’s point is that although S may say nothing about knowledge in her assertive utterance that these questions are felt by all to challenge the assertion itself and S for making them (2002, p. 263). The rationale behind this line of reasoning is that only a K-account may explain these questions as legitimate challenges to assertions since one does not explicitly state that one knows anything by asserting  $p$ .

<sup>115</sup> There are differences between Unger’s notion of ‘warrant’ and Williamson’s technical term, so when we are discussing Unger I will for the sake of clarity mark Unger’s notion with a star (\*). Unger’s account differs importantly from Williamson’s, but that has little bearing on the main points in this section. According to Unger, S *represents* herself as knowing  $p$  by asserting it (2002, p. 261). Williamson takes Unger’s idea to be upheld and subsumed under his own account, since we can explain Unger’s thesis by the general principle that whenever one does something for which authority of some kind is required one will represent oneself as having that authority. In case of assertions an epistemic authority is needed, and one therefore represents oneself as having that authority by asserting (Williamson 2000a, p. 252n6).

governed by a K-rule and one intentionally asserts without warrant\*, one intentionally conveys a false impression. Hence the aggressive question is a serious accusation, as it calls into question a warrant\* that will be on a par with lying if it is intentionally flouted. Interestingly, a response is to make an “assertive retreat” and say that what one meant was that one *merely believed p*. The sense in which that *is* a withdrawal can be explained on the K-account, since C can know that she believes *p* without knowing *p* (2002, p. 264).

#### 4.7 Moore’s paradox

An advantage of the K-account is its ability to provide a neat solution to Moore’s paradox.<sup>116</sup> Let us commence our discussion by considering the following sentences:

- (M1) It’s raining, but I *don’t believe* that it is.
- (M2) It’s raining, but I *don’t know* that it is.<sup>117</sup>

The sense in which M<sup>1</sup> and M<sup>2</sup> are paradoxical has been disputed, since the states of affairs described are not inconsistent and we could say things like, “*p*, but S does *not know/believe* that *p*” (Hintikka 1962, p. 50).<sup>118</sup> Hence there can be circumstances where the truth-conditions of M<sup>1</sup> and M<sup>2</sup> are satisfied. Yet, *any* utterance of them *sounds as* though they say something inconsistent (Unger 2002, p. 256). Why? Interestingly, we may alter them to make felicitous utterances that nevertheless describe the same state of affairs. For instance, by replacing the indexical with an anaphoric or demonstrative use of ‘she’ (M<sup>3</sup>), or altering tense (M<sup>4</sup>):

- (M3) It’s raining, but *she* does not know/believe that it is.
- (M4) It was raining, but I *didn’t* know/believe it.<sup>119</sup>

The K-account offers a neat solution to the paradox since it can explain it as unwarranted<sub>ass</sub> assertibility: We cannot assert, “*p* and I don’t know that *p*”, because anyone who makes an assertion *p* conveys that they know *p*. In case of M<sup>2</sup> one conveys the following:

- (M5) I know (It is raining, and I *don’t know* it is raining).

---

<sup>116</sup> The paradox is named after G. E. Moore, who formulated the paradox as a problem of asserting the conjunction “*p* & I don’t believe *p*” (1966, p. 78). Early work on the paradox did not always conceive it as self-contradictory since it is possible to assert *p* when S does not believe it (Hintikka 1962, p. 50). A liar asserting *p* with the intention of deceiving someone would constitute just such a circumstance. Note that Moore himself recognized the possibility of rephrasing his paradox with ‘know’ replacing ‘believe’ (Moore 1962a, p. 277).

<sup>117</sup> See Unger (2002, p. 258).

<sup>118</sup> Hintikka says that, “most people would probably admit there is something *logically* very queer about (8)” (1962, p. 50). ((8) is Hintikka’s formulation of Moore’s paradox). Williamson formulates the same puzzle when he says that, “Something is wrong with any assertion of the form ‘A, and I do not know that A’, even though such assertions would often be true if made” (2000a, p. 253).

<sup>119</sup> See Hintikka (1962, p. 50-1) and Unger (2002, p. 258).

On the *principle of distribution*, knowledge distributes over conjunction, and one can deduce knowledge of  $p$  and  $q$  from knowledge of  $(p \& q)$ .<sup>120</sup> By asserting  $M^2$  S conveys knowledge of the conjunction, “It is raining & I don’t know it is raining”, which entails  $M^6$  by distribution. From  $M^6$  and factivity of knowledge with conjunction elimination we get  $M^7$ :

(M6) I know (It is raining) & I know (I don’t know it is raining).

(M7) I don’t know (It is raining).

Since  $M^7$  contradicts the first conjunct of  $M^6$ , the K-account explains the infelicity of  $M^2$  as a contradiction that follows from the norm of assertion (Williamson 2000a, p. 253). On the original formulation of Moore’s paradox ( $M^1$ ), we employ the notion of ‘believe’ rather than ‘know’. The question is, whether the K-account can accommodate this too? I shall briefly argue that it can. We have already endorsed that if S knows  $p$  then S also believes  $p$ . By asserting, “ $p$ , but I do not believe that  $p$ ”, one conveys that one knows “ $p \& \sim Bp$ ”. Apply distribution and one conveys that one knows “ $p$ ” and that one knows “I do not believe that  $p$ ”. Again, since knowledge entails belief, the first conjunct is believed and hence one conveys, “I believe  $p$ ”, which contradicts the last conjunct and explains the paradox in  $M^1$  (2000a, p. 254).

Moore’s paradox can also demonstrate the insufficiency of a Gricean-inspired K-account of assertion: On that account the knowledge-requirement is explained as constituting a *conversational implicature* following assertions. The marker of a conversational implicature is that it is *cancellable* (either explicitly or by contextual markers) (Grice 1989, p. 39, 44).<sup>121</sup> Such an account *cannot* explain why an utterance, “ $p$ , but I do not know  $p$ ”, is wrong since that utterance explicitly cancels the implicature to know by its second conjunct (Williamson 2000a, p. 254). Moore’s paradox indicates that the requirement of knowledge is not a conversational implicature of the assertion, and provides strong evidence for the K-account.

#### 4.8 Conclusion

We have seen that observational patterns fit the account and that it provides an elegant solution to Moore’s paradox. We also found that propositions with less than probability 1 on one’s evidence are unassertable, which supports a close tie between assertion and knowledge. All this gives us strong reasons to accept the K-account of assertion.

<sup>120</sup> Hawthorne provides a good formulation of the principle saying that: “[I]f one knows that  $p$  and  $q$ , one knows  $p$  and one knows  $q$ ” (2004, p. 41). It is standard to conceive that this principle in its unrestricted formulation has intuitive appeal. Also see chapter 6.

<sup>121</sup> *Cancellability* is the mark of conversational implicatures since it is always possible to opt out of “observation” of “the Cooperative Principle” in conversation. Most conversational participants reasonably observe this principle when they want to realize the significance of communication. Implicatures are calculated on the assumption that speakers observe the principle, so by opting out one will cancel them. We may explicitly cancel it by adding a clause or we may cancel it contextually (Grice 1989, p. 29-30, 39).

## Chapter 5: Scepticism

*There is something primitive and inevitable about sceptical doubt. It runs deep in human thought.*

Colin McGinn

We have seen that Williamson could gain important explanatory advantages by employing knowledge as an unexplained explainer in the centre of his philosophical project. But there is an important sense in which that employment is a bit risky, since the notion of knowledge is also associated with the fundamental problem of epistemological scepticism. The issue of scepticism raises the question whether or not we can truly be said to know anything. In section 5.1 I commence our discussion by providing *data* and intuitions associated with the phenomenon, while in 5.2 I seek to elucidate key notions, such as the notion of a *sceptical hypothesis* and a *full sceptical account*, which will serve us in the following discussion about the structure of *classical sceptical arguments*. In 5.3 I provide two arguments for global epistemological scepticism and (near) universal ignorance, while in 5.4 I elaborate critically the sceptical position. A particularly heavy cost of scepticism is its apparent tension with intuitive normative connections between knowledge, assertion and practical reasoning, so I discuss that in 5.5. I postpone my discussion of Williamson's non-concessive response to chapter 6.

### 5.1 Scepticism and its data

Scepticism is a phenomenon that cries out for philosophical explanation (Williamson 2005c, p. 98). When competent speakers consider sceptical arguments, their confidence towards affirmative knowledge-attributions typically becomes unstable.<sup>122</sup> We can say that the sceptical arguments reliably produce certain *effects* among which are shaky knowledge-attributions. Consider the following story from Dretske (1970, p. 138):

ZEBRA: You take your son to the zoo, see several zebras, and, when questioned by your son, tell him they are zebras. Do you know they are zebras? Well, most of us would say that we did know this. We know what they look like, this is the city zoo, the enclosure where the animals reside is clearly marked "Zebras", etc. Yet, something's being a zebra implies that it is *not* a mule cleverly disguised by the zoo authorities to look like a zebra. Do you know that these animals are not mules cleverly disguised by the zoo authorities to look like zebras? Arguably not! But then could we know that they are zebras?

---

<sup>122</sup> Williamson makes the same point: "Sceptical arguments have a pull on our assent that cries out for explanation. Their power is felt immediately by a high proportion of those who are willing to listen to them carefully" (2005c, p. 98).

Typically, stories like ZEBRA raise a number of intricate questions: Does the father know that the animals are zebras? Assume that they are in fact zebras: Does the counterpossibility destroy knowledge? There is a strong intuitive pull in the direction of retracting the assertion. How are we to explain this? Why are exotic alternative hypotheses so convincing to people? Can that reveal something important about the nature of knowledge?

Let us call it the *phenomenon of scepticism* when the use of a particularly *sceptical hypothesis* (S-hypothesis) is responsible for the pull in the direction of retracting affirmative knowledge-attributions. This phenomenon is part of the wider phenomenon of unstable knowledge-attributions, which we may call the *phenomenon of knowledge-shifts*. The latter phenomenon will include the Moorean counterpart to scepticism, which is the corresponding pull in the direction of retracting ignorance-attributions. These phenomena contain *data*, which philosophers seek to explain. Epistemological data is often intuition-based, so much theorizing about knowledge is *intuition driven*. Roughly speaking, one puts epistemologically relevant intuitions into a reflective process where the output is normative claims about knowledge and epistemology. We want the best explanation so we look for an account that can make as much sense of the data as possible. One plausible constraint on the account is the *methodological principle of charity* (MPC). MPC is meant to encourage theorists to maximize rational interpretation of speakers and to interpret speakers in concordance with the *theorists'* set of beliefs.<sup>123</sup> As a natural consequence, there is a slight theoretical preference for the most charitable account vis-à-vis the phenomena of knowledge-shifts and scepticism.

Let me commence this discussion by providing a brief enumeration of data, starting with the humdrum fact that we *do* attribute knowledge. I take myself to know that I have hands, and that Pierre knows where London is. We also consider certain procedures to be reliable knowledge-producers, and hence we have a *prima facie* confidence that commits us to David Lewis' dictum, which says that, "We know a lot" (1996, p. 418). Think of this as a warrant to produce utterances like, "S knows that *p*", and thus that the beliefs expressed in such knowledge-ascriptions are sometimes true (Cappelen 2005, p. 15).

---

<sup>123</sup> The principle of charity is the methodology of interpretation advocated by Donald Davidson as a necessary ingredient in radical interpretation (1973, p. 136-7). According to him, we need to maximize agreement in order to radically interpret and understand another creature. The procedure is justified by the fact that disagreement and agreement alike are intelligible only against a background of massive agreement. A result of interpreting another with too many false beliefs, is that she comes out as neither rational, nor as having beliefs, nor as saying anything (1973, p. 137). Davidson strengthens his claim in another article, where he says that, "Charity is forced on us; whether we like it or not, if we want to understand others, we must count them right in most matters" (1974, p. 197). Jason Stanley commits himself to this methodology when he reveals that his philosophical tendency "is to preserve as much as possible of common-sense intuition" (Stanley 2005, p. v). Observe also that on this principle rationality is identified and measured by starting with what *we've* got (Brandom 2002, p. 4).

We also find knowledge desirable (Unger 2002, p. 99): It is virtuous and better to know *p* than simply believe it.<sup>124</sup> This attitude is reflected in the writings of Aristotle when he says that knowledge is a virtue and *telos* for mankind in the opening lines of his *Metaphysics*: “All human beings by nature desire to know” (Aristotle 2005a, 980a21). Williamson expresses the same affection when he takes it that “knowing matters”, and when he says that: “Even unsophisticated curiosity is a desire to *know*” (2000a, p. 31). The value of knowledge is also expressed by the importance we attach to factive mental attitudes. That importance stems from the essential matching those attitudes have with the world (Williamson 2000a, p. 40).

Despite the apparent value of knowledge, speakers often seem to lack the required dialectical weaponry to defend their knowledge base from dissolving under a sceptical assault. Hence is a result from serious consideration of sceptical arguments, that our practice becomes highly unstable. Observe that this is consistent with the impermanent nature of the effect. Typically, when people attend epistemology seminars (or are situated in some similar context), they are temporarily cautious in making knowledge-attributions. But the caution tends to *gradually loose its ground*, so that by the end of the day people return to their old practice (MacFarlane 2005, p. 211). This is not to underestimate the power of the sceptical sentiments *once they are aroused*. A sceptical argument is only interesting when its influence generates some important generalizations. Evidently, the sceptical conclusion denies knowledge, but not in a contingent manner; rather the sceptic aims to arouse the “tutored intuition” that: “Everybody is always ignorant of everything” (Unger 2002, p. 13, 94). The sceptic’s claim is radical in the sense that it denies knowledge of *any* proposition for *every* epistemic subject and at *all* times (Cappelen 2005, p. 19; Hawthorne 2004, p. 5-6; Unger 2002, p. 13, 94). The conclusion is supposed to hold both *necessarily* and *universally*, and thereby generalise in an important sense apt for normative epistemology (Vogel 2005, p. 72).

When we are confronted by a persuasive sceptical argument, we feel “caught out” or corrected as epistemic subjects (Williamson 2005a, p. 221). In other words, there is a *sense of discovery* that accompanies the argument, and we feel that we have learned something new about our epistemic situation (Cappelen 2005, p. 19). Unger, for instance, takes sceptical arguments to reveal deeper levels in our thinking previously covered by “the superficial if effective disguise of custom and intellectual lethargy” (2002, p. 1-2). That said, these sentiments have their *Moorean counterpart*. Recall our qualification of sceptical persuasiveness: Persuasion does not rule out an impermanent nature of the conviction. Our

---

<sup>124</sup> Unger qualifies: “[O]ur intuitive thought about knowledge or knowing is that it is something good, of value, which ought to be sought and prized when attained” (2002, p. 99). What’s intuitive may nonetheless prove false and replaced by the more “tutored intuition” (2002, p. 13).

next observation is that sceptical inclinations tend to be impermanent and short-lived. To speak with a devoted sceptic: “Being much like other people, I too return quickly to supposing that I know quite a fair amount. I do this almost as soon as I cease to think actively about what seems the deeper aspects of these issues” (Unger 2002, p. 9). Hence sceptical inclinations and ignorance-attributions tend to lose their grip and be unstable in a way parallel to knowledge-attributions. This goes for both sceptics and non-sceptics alike, and may explain why people are so prone to hold sceptical claims and inclinations up for scorn when they attend to different matters (Hawthorne 2004, p. 131; Williamson 2005a, p. 221).<sup>125</sup>

Speakers intuitively view affirmative knowledge-ascriptions of S (fixing content, time and place) as incompatible with corresponding ignorance-ascriptions (MacFarlane 2007, p. 17; Williamson 2005a, p. 221). Also, speakers tend to report knowledge-attributions *homophonically*, so that after a knowledge-shift they tend to *retract* earlier claims that run counter to what they think at the present. Hence we tend to make attributions in *embedded* occurrences and use the standards that are in play *now* (MacFarlane 2005, p. 202, 204).

Our last observation is that we intuitively find the sceptical conclusion to be absurd. Let me briefly comment on that: Most of us agree with the Lewisian dictum and conceive ourselves as epistemic subjects that do have some knowledge. That we do not know anything, and cannot know anything no matter how hard we try, just seems absurd, ridiculous and counterintuitive. This is a common observation among people. After all, sceptical arguments are greeted with surprise and not a shrug (MacFarlane 2005, p. 207). Does this observation rule out sceptical accounts at the outset? No. The connection between speakers’ *use* of language and its meaning is indirect, so there is no reason to take these data to rule out scepticism directly. If only the sceptic can tell us a convincing story, this would put the sceptic back in the game (MacFarlane 2005, p. 204).

This ends our enumeration, and the following list summarizes our results:

- a) We confidently make (*non-sceptic*) *knowledge attributions*.
- b) We *value* and *desire* knowledge.
- c) We tend to be (at least temporarily) *persuaded* by sceptical arguments.
- d) We take the scope of the sceptical conclusion to hold *universally* and *necessarily*.
- e) We feel caught out or corrected when the sceptical arguments make us adjust our epistemic claims.
- f) Our sceptic inclinations tend to *lose their grip* in everyday life.
- g) We take knowledge-attributions and denials regarding the same subject, time and place to *disagree*.

---

<sup>125</sup> We may label this *Hume’s Observation*, since he observed that his philosophical disposition to “withhold judgement” lost all its power outside the philosopher’s chamber: “Nature breaks the force of sceptical arguments in time, and keeps them from having any considerable influence on the understanding” (Hume 1969, p. 238).



- h) We tend to make knowledge attributions *homophonically*.
- i) We tend to find the sceptical conclusion *absurd*.

I shall take (a)-(i) to be our non-exhaustive initial set of data. The sub-set (c)-(e) and (g) constitute the basic elements in the *phenomenon of scepticism*, while the sub-set (a)-(e) and (k) are the basic elements in a *Sceptical Paradox*. By adding (f), we have the minimal set to constitute a *knowledge-shifting phenomenon*. There are tensions in these sets that call for explanation: Clearly (a) and (i) are in conflict with (c)-(e), since to take (c)-(e) seriously will shake our confidence (a) and put pressure towards revising (i). On the other hand, (a) counts as evidence against (c)-(e), while a strong conviction that (i) should *prima facie* go against (c)-(e). Some take these tensions as evidence for knowledge-shifting strategies (KSS), and to support a corresponding *variability of standards* (MacFarlane 2005, p. 200). But note that the *embedded occurrences* of ‘know’ in (g) and (h) seem to prevent us from mixing the standards; we shall see how this can be used against KSS (MacFarlane 2005, p. 202). We shall also feel the importance of observations (c) and (f), and the fact that we tend to *retract* earlier claims when they run counter to the standard that is in play at the present moment (MacFarlane 2005, p. 202). I shall take it that a full account must test its explanation against all these observations. With these brief comments I first turn to accounts that are concessive to the sceptical data (c)-(e), and non-concessive towards the Moorean data (a), (b), (f) and (i).

## 5.2 The structure of the classical argument for scepticism

The sceptic must establish the crucial intuitions (c)-(e) and (g).<sup>126</sup> Typically, she makes use of a powerful line of reasoning that employs a *sceptical hypothesis* (S-hypothesis) as one of its crucial premises (DeRose 1999, p. 1; Vogel 2005, p. 73-4). *Qua* hypothesis it explains, and its explanatory role is to make clear how one might falsely think of oneself as knowing *p* even though not-*p* is the case. Hence the content is by and large *delusive* (DeRose 1999, p. 1; Vogel 2005, p. 72). Call any further explanation of data a (full) *sceptical account*. I shall first attend to the argument and then treat sceptical accounts below.

In a *classical sceptical argument* the S-hypothesis is a proposition that seeks *minimally* to establish that one’s experience or beliefs could have been caused in a *delusory* manner. For that reason, the argument has been classified as a *Deceiver Argument* (DA) (Vogel 2005, p. 72, 75). The sceptic considers herself successful if we cannot rule out her S-hypothesis, so the point of the DA is to establish that our choice of a different hypothesis or

---

<sup>126</sup> Observe that without (g) one would be a contextualist or relativist about ‘knowledge’, and block the generalization of scepticism.

ordinary knowledge-attribution is *arbitrary* and *underdetermined*. In order to do that, the sceptic must endorse something like the following principle of underdetermination (UP):

(Underdetermination Principle (UP))      If a proposition  $q$  is a competitor to a proposition  $p$ , then one can know  $p$  only if one can non-arbitrarily reject  $q$ .<sup>127</sup>

By employing UP, the sceptic can claim victory if she shows that the choice between an S-hypothesis and some ordinary knowledge-attribution is underdetermined. In other words, the sceptic seeks to establish her S-hypothesis as a *competitor* that is incompatible with what we take ourselves to know, and yet something we cannot rule out non-arbitrarily (Vogel 2005, p. 73; DeRose 1999, p. 2).<sup>128</sup> In the relevant sense, we can rule out  $q$  arbitrarily just in case  $q$  is no less epistemically worthy of belief than  $p$  (Vogel 2005, p. 73). The function of the DA is to show that the S-hypothesis is a genuine competitor to ordinary knowledge-attributions.

Suppose I take myself to know the true proposition that, “I have hands” ( $O$ ).<sup>129</sup> A typical sceptical assault on my knowledge-claim will then entertain an *exotic proposition* or *contrast case*, like; “An omnipotent evil demon deceives me into thinking that I have hands” ( $E$ ).<sup>130</sup> I am disinclined to rule out  $E$ , but if  $E$  obtains  $O$  cannot hold. Hence we have a classical sceptical argument that challenges us to rule out an exotic contrast case  $E$  in order to know  $O$  (Unger 2002, p. 7). Philosophers have entertained many exotic propositions, such as the possibility of deceptive senses, that we might be dreaming, and the evil demon (Descartes 1998, p. 28-30). A recent story is Putnam’s brain in a vat (BIV), where one entertains the possibility that an evil neuroscientist has placed your brain in a vat and connected it to a powerful computer that provides you with the illusion that everything is perfectly normal (1981, p. 30). Also ZEBRA has the same structure, even though that story is about more humdrum circumstances. The function of the stories is the same: To show the confident knowledge-attributor how one might be deceived to believe falsely that one knows.

The sceptic, of course, *holds* that the exotic contrast case ( $E$ ) is underdetermined, and hence something we cannot rule out non-arbitrarily. We should be inclined to hold that, “S does not know that  $E$  does not obtain” ( $\sim K \sim e$ ). An effective S-hypothesis entails less exotic

<sup>127</sup> See Vogel (2005, p. 72-3).

<sup>128</sup> As Vogel puts it, the hypothesis, which the sceptic tries to establish a competitor to, can be tagged the “real world hypothesis” (RWH). The RWH is a body of beliefs that describes relations of causes and regularities in the world, and is what we usually take to be *real*. The S-hypothesis is a competitor in the sense that it tries to “take over the causal-explanatory *structure* of the RWH, but substitute within it reference to objects and properties other than the ones we take to be real” (2005, p. 75). Vogel characterises the sceptical hypothesis (“minimal sceptical hypothesis” (MSH)) accordingly: “The content of MSH is just that your experience is caused in a delusory manner, and no more: with respect to every  $Z$ , if it appears to you that  $Z$ , then something causes it to appear to you *falsely* that  $Z$ ” (2005, p. 75).

<sup>129</sup> G.E. Moore famously sought to prove the existence of an external world by making a certain gesture with his hands while saying, “Here is one hand”, and, by repeating the procedure with his other hand, he added that, “And here is another” (Moore 1962c, p. 24).

<sup>130</sup> This is a variant of Descartes’ classical sceptical hypothesis in his first meditation (Descartes 1998, p. 29-30).

ordinary propositions ( $O$ ), that we are inclined to claim knowledge of. We can formulate that as, “If S does not know that  $E$  does not obtain, then S does not know that  $O$ ” ( $\sim K\sim e \supset \sim Ko$ ). Now *modus ponens* can be put to work on our two premises securing the sceptical conclusion that, “S does not know that  $O$ ” ( $\sim Ko$ ) (DeRose 1999, p. 2).<sup>131</sup>

So far, our set-up has taken the logical relations for granted, but they are in need of clarification. According to Anthony Brueckner,  $Ko$  is a *counterpossibility* to  $\sim K\sim e$ , which means that  $Ko$  and  $\sim K\sim e$  are logically possible propositions that are logically incompatible to one another (1992, p. 43).<sup>132</sup> Hence from the affirmation of the first we can deduce the denial of the other, and vice versa. The principle of counterpossibility basically overlaps with the so-called *closure principle*, which says that knowledge is *closed* under known logical entailments. That is to say, *if* one knows that  $p$  entails  $q$ , one should be able to deduce  $q$  from knowledge of  $p$  (Brueckner 1992, p. 43). The validity of these principles is of crucial importance to classical arguments for scepticism. Unger calls it an *assumption of reasoning*, and he assumes that S must be able to apply a modest amount of reasoning, and thereby come to know propositions that follow logically upon what S already knows (2002, p. 14-16). Both principles articulate the intuitive idea that “deduction is a way of extending one’s knowledge” (Williamson 2000a, p. 117), and since they to all intents and purposes are identical I restrict myself to speak of closure. The following is a standard formulation of the Closure principle:

(Closure (C))                      Necessarily, if S knows that  $p$  and S knows that  $p$  entails  $q$ , then S knows that  $q$ .<sup>133</sup>

$C$  captures the idea that deduction can extend our knowledge, but it neglects that S is often required to *do something* in order to make the extension (Hawthorne 2004, p. 32). Let me illustrate: Mathematical relations can be extremely complex *a priori* entailments, so it seems plausible that S can know a mathematical proposition  $m$ , and that  $m$  entails  $n$ , and yet never reach knowledge of  $n$  due to the complexity of this *a priori* relation. That constitutes a counterexample to  $C$ . What S is required to do is to make the relevant inference, which indicates that  $C$  must be reformulated to yield Williamson’s principle of *intuitive closure*:

(Intuitive Closure (IC))                      Necessarily, if S knows  $p_1, \dots, p_n$  and competently deduces  $q$  and thereby comes to believe  $q$ , then S knows that  $q$ .<sup>134</sup>

<sup>131</sup> In the literature the argument is often formulated in contraposition: (1)  $\sim K\sim e$ ; and (2)  $(Ko \supset K\sim e)$ ; which *modo tollendo* gives: (3)  $\sim Ko$  (Hawthorne 2004, p. 2-6; Unger 2002, p. 7). The difference is insignificant since both schemes are valid and hinges on  $\sim K\sim e$ .

<sup>132</sup> This captures our notion of a ‘competitor’ in UP above.

<sup>133</sup> See Hawthorne (2004, p. 31). A formal formulation is provided by Dancy (1985, p. 10):  $PC^k$ :  $[Ksp \ \& \ Ks(p \supset q)] \supset Ksq$ .

<sup>134</sup> See Williamson (2000a, p. 117).

It is also intuitive that deduction may be a way of undermining one's knowledge (Stanley 2005, p. 94). Suppose that S knows that  $p_1, \dots, p_n$  at an instance  $t_1$ , and begins a lengthy deduction on the basis of which she comes to establish  $q$  at  $t_3$ . Suppose further that at  $t_2$  (sometime between  $t_1$  and  $t_3$ ) S's knowledge of  $p_1, \dots, p_n$  gets destroyed.<sup>135</sup> Intuitively, that blocks S from knowing  $q$  at  $t_3$ , which constitutes a counterexample to *IC* (Hawthorne 2004, p. 33). As a consequence, we need to formulate closure so that it rules out deductions where the required performance destroys the antecedent knowledge of premises. The following suggestions require that the antecedent knowledge be retained *throughout* the inference:

- (Multi-Premise Closure (MPC)) Necessarily, if S knows  $p_1, \dots, p_n$  competently deduces  $q$ , and thereby comes to believe  $q$ , while retaining knowledge of  $p_1, \dots, p_n$  throughout, S knows  $q$ .
- (Single-Premise Closure (SPC)) Necessarily, if S knows  $p$ , competently deduces  $q$ , and thereby comes to believe  $q$ , while retaining knowledge of  $p$  throughout, then S knows  $q$ .<sup>136</sup>

MPC and SPC are consistent with situations of self-defeating inferences from known premises.<sup>137</sup> Now let us employ the results from our above discussion, and say something about the structure of the classical argument. In order to invoke the sceptical intuitions (c)-(e), we need an effective S-hypothesis to be a counterpossibility and competitor to ordinary propositions  $O$ , which on UP will warrant the sceptical conclusion. Often one employs an exotic contrast case  $E$  that is formulated in such a way that it shows how  $O$  is underdetermined. The sceptic takes it that the choice between them is arbitrary, since we can deduce ignorance of any  $O$  by assuming MPC/SPC and the evident ignorance of  $E$ . Finally, in order to generalize properly and invoke (d), the argument must be formulated in such a way that its scope encompasses all propositions, people and instances of time.

### 5.3 Sceptical arguments

I take myself to know a lot. Not anything out of the ordinary, but still a bunch of things. For instance, I know that there is a computer in front of me, that there are cookies in the jar, and that there is a material world. Suppose a sceptic asks me the following:

BIV\*: You take yourself to know that you have a computer in front of you etc. Now tell me: Can you rule out that you are not a *recently envatted brain* (BIV\*)? Do you know that it is not the case that a

<sup>135</sup> For instance if one encounters misleading counterevidence or the inferences are self-defeating (Hawthorne 2004, p. 33).

<sup>136</sup> See Hawthorne (2004, p. 33-4).

<sup>137</sup> Note that "competent deduction" is not restricted to formally valid deductions; all one needs is that they are a priori (Hawthorne 2004, p. 35). Thus the principles are compatible with the performance of *material inferences\** from known premises to known entailments. \* I think of the kind of inferences that are good because of the content of their non-logical vocabulary (Brandom 2000, p. 85).

mad scientist has put your brain on a vat and connected it to a super-sized computer, so that all your memories are intact while a complex computer program provides you with new "impressions"? Can you rule out that all you experience is not envatted "impressions" indistinguishable from the real thing?<sup>138</sup>

I cannot rule out BIV\*, so the pressure is on me to assert ignorance of it. My putative knowledge entails the falsity of BIV\*, so by employing MPC I can come to the result that I must be ignorant of these things too. Now the sceptic expects me to retract my earlier claims, and surely we need to take the sceptic seriously.

Another sceptical argument derives its force from considerations about lotteries and *lottery propositions* (*L*). Its importance derives from the fact that it entertains mundane contrast cases rather than very exotic ones. A lottery proposition is a proposition of a sort that, even though its truth is highly likely, it is *not* something we are inclined to claim knowledge of (Hawthorne 2004, p. 5). For instance, "Lottery ticket #666 is a loser" will very likely express a truth if the lottery is big enough, but intuitively you do not know it. Now all we need is an *ordinary proposition* (*O*), which we ordinarily take ourselves to know, and which entails the truth of *L* (Hawthorne 2004, p. 5). If we can get that, we can apply SPC to generate a powerful and generalizing sceptical argument. Consider the following story:

LOTTERY: Jim is a person with moderate income who dreams of going on a safari the next summer, but who knows that with his salary this won't come true. Occasionally, he buys himself a lottery ticket in the Big Lottery. If he wins (which he knows is highly unlikely) he will go on a safari next summer.<sup>139</sup>

What the story is apt to show is that given any pair of *O*'s that entail *L*'s people could apply SPC/MPC and derive knowledge of *L* from knowledge of *O*. But that is absurd. Let me illustrate: Jim knows he won't go on a safari next summer, but that he does not go is inconsistent with him winning the Big Lottery. Hence his knowledge entails that he won't win. By applying SPC, while retaining knowledge of the fact that he won't go on a safari throughout, Jim can deduce that he won't win the Big Lottery (Hawthorne 2004, p. 2-3). But

---

<sup>138</sup> In BIV\* I stipulate that the brain is recently envatted so that we can avoid Putnam's rejection of scepticism based on externalist semantics (1981, p. 35-8). Putnam argued that it is necessarily false that we are a BIV since neither a BIV nor we can express truly that we are BIVs. The BIV cannot because her proposition will fail to express the proper content. This point is due to externalist semantics, which necessitates some kind of causal/externalist connection between the speaker and her environment in order to provide appropriate meanings to her terms. Since a BIV cannot refer to external objects, the meaning of important words is lost to her. Arguably, 'brains' and 'vats' are external objects, and thus for the BIV a sentence like, "I am a BIV", says something false. It evidently says something false when uttered by someone that is not a BIV, so the sentence is in any case false. Thus it is a necessary falsehood, and hence must its negation, ~BIV, be necessarily true (Putnam 1981, p. 37). When we stipulate a recently envatted brain we make sure that the brain maintains the appropriate meanings to a sentence like, "I am a BIV". Hence our sceptical argument is not rejected by Putnam's argument. See Brueckner (1992, p. 59n10), and also Williamson (2000a, p. 165). According to Graeme Forbes, there is a sense in which Putnam evades the genuine question raised by sceptics and the exotic contrast case BIV. As Forbes says, Putnam's reasoning "exhibits the same fallacy as in moving from "I am elsewhere now" being false in every context to the impossibility of my really, actually, possibly being elsewhere now" (Forbes 1995, p. 74n6).

<sup>139</sup> See Hawthorne (2004, p. 2-3). We could give other stories relating other lottery propositions, like "whether A will suffer from a severe heart attack next week", or, "whether my car has been stolen during the day". See Hawthorne (2004, p. 3-5).

we are unlikely to welcome this result. For one thing, it follows that other participants in equivalent epistemic positions can come to know that their tickets won't win. It is plausible that many ticket-owners are similarly situated, and so many participants can know that their ticket won't win, which seems to run counter to our beliefs about the epistemic conditions of lotteries. What's more, if Jim can deduce knowledge about his own ticket, then, by assuming we can extend knowledge by *conjunction-introduction*, nothing can stop Jim from deduce knowledge about a bunch of other tickets too (Hawthorne 2004, p. 6).<sup>140</sup> But that's crazy, and greatly overextends Jim's epistemic capacities. In fact, given enough ordinary knowledge, Jim might deduce who is going to be the winner. These perplexities provide a strong pressure on us to retract, not only Jim's original knowledge claim, but also any ordinary proposition that entails a lottery proposition. Since there is reason to believe that most ordinary propositions have such entailments, these lottery considerations comprise a powerful sceptical argument.<sup>141</sup>

There are limits to the scope of these conclusions and we can still know the following:

- (1) If the cat is on the mat, there is no evil demon deceiving me to believe falsely that the cat is on the mat.
- (2) Either I will go on a safari the next summer, or I won't win the big lottery.
- (3) It is not the case that the cat is on the mat and an evil demon deceives me to believe falsely that it is.<sup>142</sup>

Other survivors are the so-called *Cartesian propositions*, such as "I exist", "I believe that *p*", or "I am reasoning", phenomenological propositions, such as "I am in pain", knowledge of the meaning of our words and necessary truths. They all seem to avoid the sceptical arguments (Unger 2002, p. 44-5). But even though their scope does not encompass absolutely all knowledge, these arguments are sufficient to alarm the friends of knowledge.

#### 5.4 Universal ignorance

What has been said is applicable to some of the data (c)-(e), but does not constitute a full account. The final two sections will attend to the pros and cons of a *full sceptical account*. Evidently, the sceptic concedes to (c)-(e) by acknowledging (near) universal ignorance, while she regards (a) and (i) as some kind of failure since they are in disagreement (g) with (c)-(e). Also, she takes the sceptical shift in (c) to tell us something important about the nature of knowledge, while the Moorean shift in (f) does not have the same theoretical weight. A *full*

---

<sup>140</sup> Suppose there are 1 000 participants  $f_1 \dots f_{1000}$  in a lottery, and Jim knows for all of them some ordinary proposition that entails a lottery proposition. By SPC Jim can come to know that the conjunction,  $f_1$  loose &  $\dots$   $f_{1000}$  loose, is true (Hawthorne 2004, p. 6n16).

<sup>141</sup> Importantly, by varying the contents of the lottery propositions one can get different generalisations: One can for instance block knowledge of beliefs about the *future* and knowledge of *non-observational* beliefs. Debatably, one can also derive consequences for many of our *observational* beliefs by employing some standard quantum mechanics to derive a suitable lottery proposition (Hawthorne 2004, p. 3-5).

<sup>142</sup> Unger discusses similar instances of leftover knowledge (2005, p. 23-4).

*account* must also confront the intuitions (a), (f) and (i). Importantly, these data sum up three major problems to scepticism: the problem of *systematic error* (a), the problem of *variability* (f), and the problem of *common sense* (i). I will begin by articulating the most plausible sceptical account, before I confront that with the aforementioned problems. In the subsequent section I will confront scepticism with problems from assertion and practical reasoning.

All versions of scepticism take it that: (i) the semantic value<sup>143</sup> of ‘know’ is strictly invariant; and (ii) all (or nearly all) positive knowledge-ascriptions are false (Hawthorne 2004, p. 114). An expression is *invariant* only if its semantic value does not vary across contexts of use, so a *sceptical invariantist* takes it that ‘know’ express the same semantic value or property across different contexts of use (i), and rejects all (or nearly all) positive knowledge-ascriptions (ii) (Hawthorne 2004, p. 113; MacFarlane 2005, p. 205).<sup>144</sup> The property invariantly expressed by ‘know’ requires the satisfaction of a *very high* epistemic standard in order to be truly ascribed to an epistemic subject. Furthermore, the sceptic’s invariantism is *strict* or *insensitive* since it associates ‘know’ with a *fixed* high epistemic standard (MacFarlane 2005, p. 198).<sup>145</sup> The sceptical commitment to very high epistemic standard (ii) explains (c)-(e) as valid intuitions that tell us something important about the nature of knowledge, while the commitment to strict insensitive invariantism (i) can accommodate the important data (g) and (h). The sceptic disagrees with all (or near all) positive knowledge-attributions (g), and judges them homophonically to be false (h).

As we mentioned above, the *sceptical* invariantist must account for why ordinary speakers are so prone to make affirmative knowledge-ascriptions and maintain a false practice. As a response, the sceptic usually adds a third claim, (iii), to her full account. The following is a taxonomy of different sceptic positions that endorse different explanatory claims as their third endorsement: (A) *Error* theory, which takes it that ordinary people generally believe *falsely* that their knowledge-attributions are true; (B) *Exaggeration* theory, which takes it that ordinary people engage in harmless exaggeration when they attribute knowledge; and (C) The *unasserted semantic value* theory, which takes it that the semantic value of knowledge attribution is (nearly) always false, but that speakers, due to some feature

---

<sup>143</sup> Semantic value has the following three features: (i) It is what is semantically expressed by a sentence; (ii) It determines a set of truth-conditions (a function from possible worlds to truth-values); and (iii) The semantic value of a sentence is a function of the semantic values of its constituent expressions (Hawthorne 2004, p. 51-2). I shall take the semantic value of ‘know’ to be a property (Stanley 2005, p. 16).

<sup>144</sup> Contextualism about ‘know’ is the negation of invariantism. On epistemic contextualism the semantic value of ‘know’ varies across different contexts of use, and hence expresses different properties (Bach 2005, p. 51, 63; Hawthorne 2004, p. 53; MacFarlane 2005, p. 198; Ludlow 2005, p. 11, 14-5; Peter and Preyer 2005, p. 2; Stanley 2005, p. 16). See next chapter.

<sup>145</sup> This is in contrast with *sensitive* invariantism that allows the epistemic standard to shift with the subject and the circumstances of evaluation (Stanley 2005, p. 86; MacFarlane 2005, p. 198). See next chapter.

of their language, never or seldom asserts this semantic value (Hawthorne 2004, p. 114-18). Of these, I take *A* to be the most plausible. Let me briefly say why I think that.

On *B* the claim is that people are prone to harmless hyperbole and exaggeration, while on *C* speakers never really take the semantic value of what they are saying to be true. Hence it is consistent with both views that people do not believe the semantic values of their own affirmative knowledge-ascriptions (Hawthorne 2004, p. 115, 119; MacFarlane 2005, p. 206). So what goes wrong in the observations (a) and (i) is not that speakers are in error about their own linguistic practice. When people say things like, “I *know* that Hilary Rodham Clinton is running for president”, it is hyperbolic speech analogously to saying things like, “I could eat a horse!” (MacFarlane 2005, p. 206) Accordingly, speakers do not take the semantic value of such utterances to be true. The problem is only that people do *not* seem to regard the first utterance exaggerative, while the second is never taken literally. What’s more, people tend to make distinctions between hyperbolic use of words and their literal use of them; but seldom do people say things like, “I was only exaggerating my claim to know”, while it is quite suitable to say something like that about the other utterance (Hawthorne 2004, p. 120; MacFarlane 2005, p. 206). A third observation is that exaggerations disrupt inferential liaisons: People are typically unwilling to assert various ostensibly entailed claims from an exaggeration (Hawthorne 2004, p. 121). For instance, we would not embrace, “It is over three quarters of a mile long”, from an exaggerative use of, “It is a mile long”. Yet, from, “I know that *p*”, we often endorse ostensibly implied weaker claims like, “I have good reason for thinking that *p*” (Hawthorne 2004, p. 121). Those observations count against both *B* and *C*.

A suggestion, that might save these views, is that we make false knowledge-ascriptions because it would be *inconvenient* to qualify our utterances in a way that would make them true. Hence we are *speaking loosely* when we affirmatively attribute knowledge. But there is a sense in which people do not recognize themselves as speaking loosely when they attribute knowledge (MacFarlane 2005, p. 207). If one challenges loose speaking, an adequate response is to retract the loose claim and respond with something more accurate, like, “I meant that so-and-so”. But people do not respond like this when we challenge their knowledge-attributions, and so, again, observations count against *B* and *C* (2005, p. 207). Another worry is that it is very hard to see that people should experience that they learn something new from sceptical arguments if *B* or *C* are true, and thus it is hard to see how they should accommodate (e) in the above data. Finally, we may worry about whether these views posit too great a gap between semantic value and linguistic practice (Hawthorne 2004, p.



123). From where does its purported semantic value come from if knowledge-ascriptions are never used literally? I take these worries to count strongly against alternatives *B* and *C*.

An error theory (*A*) is perhaps the most plausible of the sceptical options, so in the rest of this section I shall restrict myself to *A*. On *A* people make false knowledge-ascriptions without being aware of their mistake, so a first objection is that it runs counter to our data (i). In our initial set (i) represents the commonsensical response to scepticism, so this objection is really that *A* runs counter to *common sense* (Hawthorne 2004, p. 126). That *is* a cost on a full *A*-account of scepticism, but, as Hawthorne notes, commonsense is a “double-edged sword” since a commitment to the ignorance of lottery propositions is *also* commonsense (2004, p. 126). Hence we should not overstate the importance of ordinary practice and commonsensical data, and we cannot take (i) to be a knockdown consideration against *A*.

Another worry is that a full *A*-account should be able to explain why people are so prone to make systematic errors, and hence confront data (a) in our initial set-up (MacFarlane 2005, p. 211). I will grant that such a story may be given since it is plausible that one could employ the *S*-hypothesis for that purpose. Perhaps one could assert that ordinary speakers have not yet considered scepticism in a proper manner, but that once they do people will change their ways and retract affirmative knowledge-ascriptions in accordance with data (c)-(e). The problem is only that any such story will run against observation (f) and the problem of variability (Hawthorne 2004, p. 131; MacFarlane 2005, p. 213). Variability and the observation of knowledge-shifts is a problem for any kind of *strict invariantism* (MacFarlane 2005, p. 205). For the sceptic that problem is to explain the disposition and willingness in ordinary speakers to let oneself lapse into a faulty practice of making affirmative knowledge-ascriptions. These considerations require a fancy story from the sceptic, which is very seldom told. I take the abovementioned considerations to count against *A*, but since none of them conclusively rules out this variant of scepticism we must proceed to some real trouble spots.

### **5.5 Where ignorance enjoins silence as well as nihilism**

A serious worry is that scepticism seems to be in conflict with normative connections between knowledge, assertion, and practical reasoning. Let us begin with assertion: Universal ignorance or error-scepticism seems to be at odds with the knowledge account of assertion (*K*-account). Recall that the *K*-account strictly forbids one to assert *p* if one does not know it. We have already found that conception intuitive, so what is problematic for sceptics is that universal ignorance commits them to the claim that no assertions are warranted<sub>ass</sub>. For one thing, that makes error-scepticism a *weakly self-defeating* position, since no sceptic can assert

it without representing herself somewhat paradoxically as knowing the correctness of universal ignorance (Hawthorne, 2004, p. 133; Unger 2002, p. 265).<sup>146</sup> Hence our first objection is that universal ignorance enjoins silence *for* the sceptic.

That is a trouble spot for the sceptic, but it is not much of a problem for the position of error-scepticism itself, and we should make sure that we also refute the position (Hawthorne 2004, p. 133).<sup>147</sup> The *real* problem of assertion is that universal ignorance renders it necessary that *no* assertion can be warranted<sub>ass</sub>, since nobody can assert anything without giving the wrong impression. Hence ignorance threatens to enjoin silence for all of us. That is in tension with established practice, and counts against any error-scepticism that retains the intuitive connection between knowledge and assertion (Hawthorne 2004, p. 133). Of course, there is the option to abandon that connection. Perhaps one could say that assertion is normatively connected to be as close to knowledge as the context demands, or something like that (Hawthorne 2004, p. 134). But however that story unfolds, it still counts against error-scepticism that it strains the intuitive normative connection between knowledge and assertion.

Let us also attend to the second trouble spot for scepticism: There is an intuitive link between knowledge and practical reasoning, since one is prohibited from using a premise in one's deliberations about how to act if one does *not know* that premise. In other words, knowledge is the norm of practical reasoning (Hawthorne 2004, p. 29). On error-scepticism nobody knows *any* proposition, so evidently one cannot know the premises that are required for practical reasoning. Hence error-scepticism views all those deliberations improper, and consequently there is a threat of nihilism added to error-scepticism (2004, p. 134-5).

## 5.6 Conclusion

Let me summarize the pros and cons of a full account of error-scepticism: It respects and employs closure (SPC and/or MPC), and it combines those principles with the datum that we do not know the lottery propositions, yielding the sceptical conclusion. This respects data (c)-(e). What's more, an error-sceptic sticks to his claim about universal ignorance, and that easily accommodates (g) and (h). The costs are significant: It fits disastrously with data (a) and (i), and it has a hard time explaining (f) since it is hard to explain variability and shifts in Moorean direction. Error-scepticism is also at odds with intuitive normative connections between knowledge, assertion and practical reasoning. The position is not entirely disastrous, but it is not *very* attractive either. Can we do better? The next chapter will decide.

---

<sup>146</sup> As Unger puts it, "a sceptic who would communicate is in an unfortunate situation" (2002, p. 271).

<sup>147</sup> This distinction matters since, "Certain arguments that might be leveled against the skeptic – in particular arguments to the effect that the error skeptic holds self-defeating commitments – may not cut any ice against the thesis of error skepticism" (Hawthorne 2004, p. 125-6).

## Chapter 6: Anti-Scepticism

*Scepticism is a disease in which healthy mental processes run pathologically unchecked.*

Timothy Williamson

*A reasonable man does not have certain doubts.*

Ludwig Wittgenstein

Here are three natural ways of explaining the phenomena of knowledge shifts and scepticism: (i) embracing scepticism; (ii) attempt some *verbal hocus pocus*; or (iii) abandon the plausible sounding closure principle (Hawthorne 2005, p. 27).<sup>148</sup> We have already dealt with (i) and found it unattractive, so the question is: Are (ii) or (iii) any better? We will spend the final chapter addressing that issue. We start by confronting scepticism with some basic assumptions on evidence and rationality in section 6.1, before defending epistemic closure and rejecting option (iii) in section 6.2. That leaves us with option (ii), where there are three obvious choices: (A) *Contextualism*, where one treats ‘know’ as context sensitive; (B) *(Subject) Sensitive Invariantism*, where one seeks to mimic A without embracing its counterintuitive consequences; and (C) *Anti-Sceptical Insensitive Invariantism*, which seeks to explain the shifts by invoking extra-linguistic mechanisms.<sup>149</sup> I discuss A and B in sections 6.3 and 6.4, before I turn to C in section 6.5. Williamson opts for C, so we may employ resources from his account of knowledge and use that against the sceptic. Arguably, all solutions carry undeniable costs, so preference is a question of assessing their pros and cons. In section 6.6 I discuss whether Williamson’s candidate for an extra-linguistic mechanism echoes empirically, before in section 6.7 I examine the intuitive connection between knowledge and practical reasoning in order to provide some conclusions to our discussion.

### 6.1 Rational management of one’s evidence

*Some* of the abovementioned sceptical deliberations hinge crucially on debatable assumptions about the epistemic agent’s (S) access to her own evidence. Rational thinkers must respect their evidence, but does it require complete access to one’s evidence? Is it necessary for S to always be able to know what her evidence is? And how can this be reconciled with the fact that our minds are characteristically anti-luminous?

Typically sceptics compare a *good* with a *bad case* in order to reject knowledge of any *p* by appeal to UP. If S believes *p*, she cannot rule out being situated in a bad case where S

---

<sup>148</sup> The list of alternative options is non-exhaustive. We could add the Moorean response, i.e. where one bites the bullet and infer knowledge of the negation of scepticism from knowledge of ordinary propositions (Hawthorne 2004, p. 142-3). I will not discuss this position.

<sup>149</sup> This list is also non-exhaustive since I omit discussing Relativism (R<sup>l</sup>), or the position that explains knowledge-shifts as determined by the contexts of evaluation. See MacFarlane (2005; 2007).

believes  $p$  falsely, so by UP we must reject knowledge of  $p$  (Williamson 2000a, p. 165).<sup>150</sup> The conclusion follows because things appear to be *exactly the same* in both cases. If  $S$  does not know which case she is situated in,  $S$  cannot tell whether or not  $p$  is really true. According to the sceptic, there is *symmetry of evidence* between the two cases, and so  $S$  has exactly the *same evidence* in both cases (2000a, p. 164). Since  $S$ 's evidence in the good case is not inconsistent with being situated in a bad case, the assumption of symmetry provides a powerful argument in favour of the *sceptical hypothesis*. Here's why: *If*  $S$  is in a good case and cannot rule out being in a bad one, and  $S$ 's evidence is the same in both situations, *then*  $S$ 's *true* belief that  $p$  is due to *mere epistemic luck*. This is evident by the fact that  $S$  could have had the exact same evidence in a case where  $p$  is false (2000a, p. 166). We have already accepted that knowledge is incompatible with epistemic luck, so  $S$  cannot know  $p$  in the good case (Steup 2006). Note that on evidential symmetry  $S$ 's evidence in a bad case is consistent with  $S$  being situated in a good one, since symmetry works in both directions (Williamson 2005a, p. 165-6). Consequently, we cannot tell which case we are situated in based on our evidence. When  $S$  is required to respect her evidence, she is expected *not* to draw different conclusions based on identical sets of evidence. She believes  $p$  falsely in the bad case and therefore cannot know  $p$ , yet she has the exact same set of evidence in a good case, and hence must her evidence be insufficient for knowing (Williamson 2000a, p. 169).

*Sameness of evidence* is a crucial premise in these sceptical arguments, and needs to be justified. Consider the following *reductio ad absurdum*: Suppose (with the aim of deriving a contradiction) that the evidential situation is *different* in good and bad cases. It follows that  $S$  can deduce that she is *not* in a good case when she is in a bad one by comparing evidence (Williamson 2000a, p. 169-70). But that is absurd. Even anti-sceptics recognize that when she is in a bad case  $S$  cannot know that she is not in a good one (2000a, p. 170). By negation-introduction, we can assert that one has the *same* evidence. Let us examine this argument more closely: The *reductio* depends on epistemic access to one's evidence in a bad case, so we must start with the premise that one knows what one's evidence is ( $S_i$ ) (2000a, p. 170-1):

( $S_i$ ) For any appropriate property  $\pi$ , in any case in which one's evidence has  $\pi$ , one knows it.

In order to make the comparison must  $S$  know in the bad case what her evidence *would be* if she were situated in a good case ( $S_{ii}$ ). A basic assumption is also that if  $S$  is in a bad case, she cannot deduce that she is not in a good case after all ( $S_{iii}$ ) (2000a, p. 171).

---

<sup>150</sup> UP: If a proposition  $q$  is a competitor to a proposition  $p$ , then one can know  $p$  only if one can non-arbitrarily reject  $q$ . See chapter 5.

- (Sii) For any appropriate property<sup>151</sup>  $\pi$ , if in the good case one's evidence lacks  $\pi$ , then in the bad case one knows that in the good case one's evidence lacks  $\pi$ .
- (Siii) It is consistent with what one knows in the bad case that one is in the good case.

The next two premises jointly constitute the assumption of the *reductio* (2000a, p. 172):

- (Siv) In the bad case one's evidence has  $\pi$ .
- (Sv) In the good case one's evidence lacks  $\pi$ .

Now (Sii) and (Sv) jointly entail (Svi), while (Si) and (Siv) entail (Svii) (2000a, p. 173):

- (Svi) In the bad case one knows that in the good case one's evidence lacks  $\pi$ .
- (Svii) In the bad case one knows that one's evidence has  $\pi$ .

The content-clauses of what one knows in (Svi) and (Svii) jointly entail that, "One is not in the good case". By MPC<sup>152</sup> S can expand knowledge to yield (Sviii) (2000a, p. 172):

- (Sviii) It is inconsistent with what one knows in the bad case that one is in the good case.

This result (resting on (Si), (Sii), (Siv) and (Sv)) clearly contradicts (Siii), so by assuming (Si)-(Siv) we can deny (Sv) by *reductio ad absurdum*. That yields (Six), which can be conditionalized on (Siv) so that we get (Sx) (Williamson 2000a, p. 172):

- (Six) In the good case one's evidence has  $\pi$ .
- (Sx) If in the bad case one's evidence has  $\pi$ , then in the good case one's evidence has  $\pi$ .

We can now run an analogous argument replacing " $\pi$ " with "non- $\pi$ ", and establish (Sxi). Contraposition<sup>153</sup> on (Sxi) provides the converse of (Sx), thus enabling us to make the generalized claim (Sxii), which articulates evidential symmetry (2000a, p. 172):

- (Sxi) If in the bad case one's evidence has not- $\pi$ , then in the good case one's evidence has not- $\pi$ .
- (Sxii) One's evidence in the good case has same appropriate properties as one's evidence in the bad case.

---

<sup>151</sup> An "appropriate property" is a specification of content in evidence. I will grant, for the sake of argument, that one can know that the evidence has this property under some canonical specification. We shall also assume that if a property is appropriate, then so is its complement (2000a, p. 170). That assumption warrants talk about lack of appropriate property  $\pi$  as well as  $\pi$  itself.

<sup>152</sup> Multi-premise Closure (MPC) allows S to deduce and come to know the content of known entailment(s) from a multitude of known premises (Hawthorne 2004, p. 33). See previous chapter.

<sup>153</sup> The contraposition of (11) is (11\*): If in good case S' evidence has not non- $\pi$ , then in bad case S' evidence has not non- $\pi$ . By removing double negation we get (11\*\*): If in good case S' evidence has  $\pi$ , then in bad case S' evidence has  $\pi$ . And (11\*\*) is the converse of (10).

Comment on the result is mandatory: The argument is formally valid, so *if* its premises are sound, the conclusion ( $S_{xii}$ ) should impose severe restrictions on the nature of evidence. By accepting it much is conceded to the sceptic, as for instance, that evidence drives towards the purely phenomenal (or towards “those conditions, whatever they are, which rational subjects can know themselves to be in whenever they are in them”) (Williamson 2000a, p. 173). The argument deprives us of any pretension to know  $p$  on one’s total evidence, since an identical set of evidence in a bad case is basis for a false belief  $p$ . After all, this is what makes the argument so apt for sceptics, since there is always a corresponding bad case to a good case (2000a, p. 174). What’s more, the argument is not vulnerable to objections that employ the distinction between relevant and irrelevant alternatives. That follows since the argument hinges on claim ( $S_{iii}$ ), which says that it is consistent with what one knows *in the bad case* that one is in a good (2000a, p. 174). Thus appeals to the irrelevance of scepticism and bad cases for agents in good cases will not be sufficient to block this argument.

Let me briefly expand on that: According to Fred Dretske, we can successfully block the infelicitous inferences from knowledge of an ordinary proposition to knowledge of the negation of the sceptical hypotheses, since one’s evidence need only rule out *evidentially relevant* alternatives in order to have knowledge of the first proposition (Dretske 2005, p. 19). Arguably, a bad case is not evidentially relevant for S in a good case, so what might go wrong in bad cases is not relevant for S when she is situated in a good case. But our above sceptical argument is consistent with Dretske’s distinction, since it concentrates on the epistemic status and deliberations of the agent in a bad case. Thus Dretske’s requirement is insufficient. Perhaps one could turn the point and deny the relevance of good cases for the agent in a bad case, but that move would only serve to obscure the debate. When one is situated in a bad case one thinks of oneself as being situated in a good, which makes the good case relevant.

One natural response is to bring the explanatory resources of Williamson’s account of knowledge into play in order to secure an anti-sceptical counterclaim.<sup>154</sup> Recall that one’s total evidence is equated with one’s total knowledge ( $E=K$ ) (2000a, p. 180, 185).<sup>155</sup> As a consequence, S cannot know  $p$  in a good case where S’s evidence is equated with the evidence in a bad case. Here’s why: If  $p$  is not part of S’s total knowledge in the bad case and  $E=K$ , then  $p$  is not part of S’s total evidence due to the *Indiscernibility of Identicals*. On the sameness of evidence premise (i.e.  $e$  in Bad =  $e$  in Good), and with another application of

---

<sup>154</sup> An analogous argument is also given in Williamson (2000b). I shall keep the following discussion to his argument in chapter 8 (2000a).

<sup>155</sup> See chapter 3. The “E = K thesis” says that there is no evidence that is not part of S’ total knowledge (Williamson 2000a, p. 185).

Indiscernibility of Identicals, we can infer that S does not know  $p$  in the good case.<sup>156</sup> In other words, that S does not know  $p$  in the good case follows trivially on the sceptic's assumption of sameness of evidence. When the sceptic equates the evidence in the two cases, that does not establish the claim that one does not know  $p$  in the good case no more than the fact that one can infer any  $p$  from itself establishes  $p$  (Williamson 2000a, p. 180).

The problem with this reply is that it hinges on the equation  $E = K$ , and a sceptic will probably reject that equation. Let us proceed to a different objection: One option is to dispute (S<sub>i</sub>). Perhaps one can parallel the sceptic's argument and produce a line of reasoning that warrants obvious falsehoods from the assumption (S<sub>i</sub>)? Again, by negation-introduction one can deny the assumption and the sceptic's argument will be a non-starter. Let us run through the argument: Assume (S<sub>i</sub>) with the aim of deducing a contradiction. Let  $t_0, t_1, \dots, t_n$  be a long sequence of times at one-millisecond intervals, and assume that across the time-span one's evidence changes gradually so that appropriate properties of one's evidence are different by the end of the interval.<sup>157</sup> By paralleling premise (S<sub>ii</sub>) we get (2<sub>i</sub>) (Williamson 2000a, p. 175):

(2i) For any appropriate property  $\pi$ , if in  $\alpha_{i-1}$  one's evidence lacks  $\pi$ , then in  $\alpha_i$  one knows that in  $\alpha_{i-1}$  one's evidence lacks  $\pi$ .<sup>158</sup>

Now assume both limited power of discrimination (3<sub>i</sub>) and (4<sub>i</sub>), and then assume (5<sub>i</sub>) for a *reductio ad absurdum* (2000a, p. 176):

(3i) It is consistent with what one knows in  $\alpha_i$  that one is in  $\alpha_{i-1}$ .

(4i) In  $\alpha_i$  one's evidence has  $\pi$ .

(5i) In  $\alpha_{i-1}$  one's evidence lacks  $\pi$ .

Now premises (2<sub>i</sub>) and (5<sub>i</sub>) entail (6<sub>i</sub>) while (S<sub>i</sub>) and (4<sub>i</sub>) entail (7<sub>i</sub>) (2000a, p. 176):

(6i) In  $\alpha_i$  one knows that in  $\alpha_{i-1}$  one's evidence has  $\pi$ .

(7i) In  $\alpha_i$  one knows that one's evidence has  $\pi$ .

From the content-clauses (6<sub>i</sub>) and (7<sub>i</sub>) we may deduce, "One is not in  $\alpha_{i-1}$ ". By MPC and knowledge of both (6<sub>i</sub>) and (7<sub>i</sub>) we may deduce (8<sub>i</sub>), which contradicts (3<sub>i</sub>) on assumptions

<sup>156</sup> The *Indiscernibility of Identicals* reads:  $c = d \supset (X)(Xc \equiv Xd)$ . This is the converse of the *Identity of Indiscernibles*, which reads:  $(X)(Xc \equiv Xd) \supset c = d$ . *Leibniz' Law* joins these in the following biconditional to define identity:  $c = d \equiv (X)(Xc \equiv Xd)$  (Boolos, Burgess and Jeffrey 2003, p. 280). In our case the epistemic set  $\Gamma$  is identical to  $\Gamma^*$  ( $\Gamma = \Gamma^*$ ), so by the Indiscernibility of Identicals it follows that all appropriate properties  $\pi$  on  $\Gamma$  is a property on  $\Gamma^*$  ( $\pi\Gamma \equiv \pi\Gamma^*$ ). Such a property is "p being a part of  $\Gamma$ ", so since  $p$  is a part of  $\Gamma$  it is also a part of  $\Gamma^*$ .

<sup>157</sup> E.g. at  $t_0$  one's evidence has property  $\pi$  while lacking it at  $t_n$ .

<sup>158</sup> The use of " $\alpha_{i-1}$ " is an abbreviation of a case  $\alpha$  that obtains at an instance of time  $t_{i-1}$ . Again, we restrict ' $\pi$ ' to appropriate properties.

(S<sub>i</sub>), (2<sub>i</sub>), (4<sub>i</sub>), and (5<sub>i</sub>) (Williamson 2000a, p. 176). By assuming (S<sub>i</sub>) and (2<sub>i</sub>)-(4<sub>i</sub>), we can deny (5<sub>i</sub>) by a *reductio*, yielding (9<sub>i</sub>):

(8i) It is inconsistent with what one knows in  $\alpha_t$  that one is in  $\alpha_{t-1}$ .

(9i) In  $\alpha_{t-1}$  one's evidence has  $\pi$ .

We can conditionalize (9<sub>i</sub>) on (4<sub>i</sub>), yielding (10<sub>i</sub>). An analogous argument can now be run again with “not- $\pi$ ” replacing “ $\pi$ ”, which leaves us with (11<sub>i</sub>). Now we can generalize and secure (12<sub>i</sub>), since the contraposition of (11<sub>i</sub>)<sup>159</sup> yields the converse of (10<sub>i</sub>) (2000a, p. 176):

(10i) If in  $\alpha_t$  one's evidence has  $\pi$ , then in  $\alpha_{t-1}$  one's evidence has  $\pi$ .

(11i) If in  $\alpha_t$  one's evidence has not- $\pi$ , then in  $\alpha_{t-1}$  one's evidence has not- $\pi$ .

(12i) One's evidence in  $\alpha_{t-1}$  has the same appropriate properties as one's evidence in  $\alpha_t$ .

The problem with (12<sub>i</sub>) is that the relation between cases is transitive, so repeated application establish that no properties of one's evidence change during the time-span (13<sub>i</sub>):

(13i) One's evidence in  $\alpha_0$  has the same appropriate properties as one's evidence in  $\alpha_n$ .

But (13<sub>i</sub>) is clearly false and runs counter to our assumption that one's evidence at the end of the process *is* different from one's evidence at the beginning. Thus (S<sub>i</sub>) is shown to be false, and as a consequence we can conclude that the sceptical argument is *not sound* (2000a, p. 177).

What we do in the above strategy is to sacrifice some self-knowledge in order to win knowledge of the external world (Williamson 2000a, p. 164). Williamson's point is that one does not always know the appropriate properties of one's evidence. That claim is on a par with the anti-luminous nature of our minds, and it follows as a consequence of equating one's total evidence with one's total knowledge.<sup>160</sup> Since knowledge is a mental state, the anti-luminosity of mind will sometimes hinder one from knowing what one's evidence is. Paradoxically, the sceptic must argue for an all-encompassing self-knowledge in order to combat this anti-sceptical result. An agent, who is situated in a bad case, has evidence that *looks* similar to one's evidence in a good case. But the appearance is only illusory, and the

<sup>159</sup> Contraposition of (11i) yields (11i\*): If in  $\alpha_{t-1}$  one's evidence has not not- $\pi$ , then in  $\alpha_t$  one's evidence has not not- $\pi$ . By removing double negation we get (11i\*\*) which is the converse of (10i): If in  $\alpha_{t-1}$  one's evidence has  $\pi$ , then in  $\alpha_t$  one's evidence has  $\pi$ .

<sup>160</sup> See chapters 2 and 3.



truth is that the badly situated agent has a radically impoverished mind that is not aware of its own impoverishment (Williamson 2000a, p. 177).

Let me comment on the result: Our anti-sceptical argument works against the classical arguments for scepticism, since those arguments hinge on a comparison between bad and good cases. That is an important result, but since the lottery argument makes no use of comparison it is no threat to the kind of scepticism that is motivated by lottery-considerations. The rest of this chapter will discuss different responds to that kind of scepticism.

## 6.2 Defending closure

Some philosophers<sup>161</sup> have argued against scepticism by denying the validity of the closure principle. Doing this will block most sceptical arguments, including the above lottery-considerations, since an inference from ignorance of a sceptical hypothesis (or lottery proposition) to ignorance of ordinary propositions requires closure. The *closure* principle validates an inference from known premises and a known entailment of those premises, to knowledge of the entailment. In this sense it express the compelling thought that deduction is a way of extending one's knowledge (Hawthorne 2005, p. 29; Stanley 2005, p. 18n1; Williamson 2000a, p. 117). Denying closure is not to deny that one can never extend one's knowledge by deduction, rather one reject it as a matter of principle from *any* known premises to *any* conclusion known to be entailed by those premises (Dretske 2005, p. 17). Consequently, the task for closure-sceptics is to provide a *suitable restriction* on closure and not a flat-out denial (Dretske 2005, p. 19).

Dretske's original idea was to erect that restriction from a distinction between *sensitive* and *insensitive* beliefs, and he took it that only sensitive beliefs could be knowledge on one's evidence. Hence Dretske restricted the closure principle to be valid only among beliefs that are sensitive to the same evidence. That distinction proved too crude since it ruled out knowledge like, "I believe a animal is in front of me", from observing a dog in front of me (Hawthorne 2004, p. 42).<sup>162</sup> Dretske's latest proposal is to delineate between *heavyweight* (HW) and *lightweight* (LW) implications, and allow validity of the closure principle between propositions of the latter type while rejecting inferences from LWs to HWs (Dretske 2005, p. 16).<sup>163</sup> The HWs are propositions that are out of range for ordinary perception (or the exercise of reason alone) in the sense that we cannot see, hear, smell, or feel that they are true (Dretske

---

<sup>161</sup> Nozick (1981) and Dretske (1970, 2005).

<sup>162</sup> The belief is insensitive to the evidence since one could believe truly that, "I have an animal in front of me", without observing a dog. For suppose that behind the dog there is a duck facsimile. Then absent the observation of a dog I would still believe, "I have an animal in front of me", thus rendering the belief insensitive to the evidence (Hawthorne 2004, p. 42).

<sup>163</sup> For the sake of argument I grant the distinction, but see Hawthorne (2004, p. 42-6) for some good counterexamples.

2005, p. 20; Hawthorne 2004, p. 42). For instance, we can see that there are cookies in the jar, but not that there is a material world (Dretske 2005, p. 19-20).

Denying closure seems extremely questionable: For one thing it goes against *pre-theoretic* intuitions (Williamson 2000a, p. 118).<sup>164</sup> A related worry is that it would attribute systematic error to competent speakers since *they* perform such inferences about the beliefs of other people.<sup>165</sup> If Dretske is right, we are prone to forming false beliefs about what other people know and believe by applying a non-valid closure principle (Hawthorne 2005, p. 31). This last objection needs qualification: To count against Dretske the objection must show that speakers perform such inferences to attribute beliefs of *heavyweight implications* on others. But it seems implausible that ordinary speakers should entertain heavyweight beliefs in the first place, and so the claim that they should attribute it on others on inferential ground seems strained.<sup>166</sup> Fortunately, it is possible to provide stronger evidence against closure-scepticism: The following principles seem extremely compelling (Hawthorne 2005, p. 31):

(Equation Principle (EP)) If one knows *a priori* (with certainty) that *p* is equivalent to *q* and one knows *p*, and one competently deduces *q* from *p* (retaining one's knowledge that *p*), one knows *q*.

(Distribution (D)) If one knows the conjunction of *p* and *q*, then as long as one is able to deduce *p*, one is in a position to know that *p* (and the same goes for *q*).

Dretske's sensitivity-requirement cannot block *EP* since it is absurd to assert that some evidence *R* to which *p* is sensitive is not at the same time evidence for *q*; after all *q* is equivalent to *p*.<sup>167</sup> But by granting *EP* the closure-sceptic is forced to deny *D*. For suppose *S* has evidence *R* and knows (1) (Hawthorne 2005, p. 32): Since (2) is equivalent to (1) we can now come to know (2) by *EP*. Apply *D* on (2) and (3) results:

- (1) G is full of wine.
- (2) G is full of wine and it is not the case that G is full of non-wine that is coloured like wine.
- (3) It is not the case that G is full of non-wine that is coloured like wine.

But (3) *is* heavyweight and the only way we can deny this implication while retaining *EP* is by rejecting *D*. Another compelling principle is (Hawthorne 2004, p. 39):

---

<sup>164</sup> Williamson says, "We should in any case be very reluctant to reject intuitive closure, for it *is* intuitive" (2000a, p. 118).

<sup>165</sup> E.g. if A knows that B believes both that, "I am taller than A, and A is taller than C", and that, "If I am taller than A and A is taller than C, then I am taller than C"; then A is prone to attribute the belief that, "B is taller than C", to B. Scott Soames observes the same tendency when he says that, "often we credit a person with having established that Q on the grounds that the person has explicitly established something else that has the proposition that Q as a trivial consequence" (1999, p. 26).

<sup>166</sup> E.g. the case must show something like: Let us say A forms the belief that B believes there is a material world on basis of B knowing both, "There are cookie in the jar", and, "If there are cookies in the jar there exists a material world".

<sup>167</sup> Sensitive can be cashed out by saying that: *p* is sensitive to *R* just in case, if *p* were not the case, *R* would not be the case (Hawthorne 2005, p. 31).

(Addition Closure (AC)) Necessarily, if S knows  $p$  and competently deduces ( $p$  or  $q$ ) from  $p$ , thereby coming to believe ( $p$  or  $q$ ), while retaining knowledge of  $p$  throughout, then S knows ( $p$  or  $q$ ).

Again, we can force the closure-sceptic to reject a compelling principle. Suppose I know (4): Then by AC we can deduce (5), which is equivalent to (6) (Hawthorne 2004, p. 40):

- (4) I have a hand.
- (5) Either I have a hand or I am not a brain in a vat.
- (6) It is not the case that I lack hands and am a brain in a vat.

But (6) is the antithesis to scepticism and by all standards heavyweight, so we can force the denial of AC on Dretske. Can it get any worse? Arguably it can, because the denial of closure interacts disastrously with the knowledge account of assertion (Hawthorne 2004, p. 39). The fundamental norm of assertion is that one must not assert  $p$  without knowledge thereof (Hawthorne 2005, p. 32; Williamson 2000, p. 243).<sup>168</sup> Consider the following case: S knows that the animal in the cage is a zebra (Z), and that if the animal in the cage is a zebra then it is not a cleverly disguised mule (if Z then not-M). Since she knows that S can affirm the following questions without breaking the rule of assertion: “Z?” And, “If Z then not-M?” Yet, she cannot affirm the question: “Not-M?” Since if she did, S would assert, “Not-M”, and be liable to criticism since she did not know it.<sup>169</sup>

Dretske accept these points, but he thinks that the alternatives are even more theoretically expensive (2005, p. 43). As he says: “Philosophy is a business where one learns to live with spindly brown grass in one’s own yard because neighboring yards are in even worse shape. From where I sit, the grass on the other side of the fence looks worse than mine” (Dretske 2005, p. 43). The closure-denying yard *is* in a bad shape, so I think we are required to cross the fence and check out this neighbourhood for ourselves.

### 6.3 Knowledge-shifting strategies

It has become fashionable to advocate partly concessive approaches to scepticism, where one acknowledges a certain bite to their arguments, while seeking to tame their most worst implications. In particular, philosophers have sought to rescue ordinary knowledge-claims by denying the relevance of sceptical epistemology for these claims. Perhaps philosophy itself is to blame for scepticism? Lewis even suggests that, “epistemology is the

---

<sup>168</sup> See chapter 4.

<sup>169</sup> Hawthorne likens S with the Tortoise in Lewis Carroll’s story (Hawthorne 2004, p. 39; 2005, p. 32).

culprit”, an extraordinary pastime that “robs us of our knowledge” (Lewis 1996, p. 420). This basic idea is caught up in different versions of *knowledge-shifting strategies* (KSS): A KSS is an explanatory effort based on semantics, and thus an instance of option (ii). Typically, it is a part of a general strategy to explain the *phenomenon of knowledge shifts* or unstable knowledge attributions. The oldest and most popular view is *epistemological contextualism* (EC<sup>k</sup>),<sup>170</sup> but it has been challenged lately by its younger sibling (*subject*)-sensitive invariantism (SSI<sup>k</sup>).<sup>171 172</sup> Williamson has recently disputed the cogency of both of these views for reasons that are highly relevant to his overall aim of refuting scepticism, and which may help us to understand better the distinction between knowledge and ignorance.

Typically philosophers take stories like the following to support KSS:

Normally, I am perfectly happy to say that I know that my car is parked in my driveway. I will say this even when I’m at work, several miles away. But if someone asks me how I know that my car has not been stolen (and driven away), I will admit that I do not know this. And then I will have to concede that I do not know that my car is in my driveway: after all, if I knew this, then I would be able to deduce, and so come to know, that it has not been stolen.<sup>173</sup>

This story includes some familiar basic elements: We have an affirmative knowledge attribution, K[S,*p*,*t*], that is rejected when S realizes that *q* (an entailment of *p*) is not known. Thus S comes to affirm, ~K[S,*p*,*t*], in accordance with closure. Unlike the phenomenon of scepticism this kind of conversation is ubiquitous in our linguistic behaviour and that alone warrants our attention. The fundamental question is how one should think of a shift where someone claims to know *p* and then later goes on to deny it (MacFarlane 2005, p. 200).

Both EC<sup>k</sup> and SSI<sup>k</sup> have their strategies to accommodate these phenomena. The guiding idea behind EC<sup>k</sup> is that knowledge-ascriptions are subtly context-dependent in a distinctly epistemological way (Lewis 1996, p. 420; Stanley 2005, p. 16). According to EC<sup>k</sup>, the context-dependency of knowledge-ascriptions is due to the context-sensitivity of the verb ‘know’, which makes it possible for a given knowledge-ascribing sentence *P* to express

---

<sup>170</sup> According to Stanley, EC<sup>k</sup> has its roots in Fred Dretske’s Relevant Alternative Theory (RAT). On RAT, a proposition is known if S believes it truly and is able to rule out relevant alternatives to the proposition at that time (Stanley 2005, p. 17). A relevant alternative is defined as “an alternative that *might have been* realized in the existing circumstances if the actual state of affairs had not materialized” (Dretske 1970, p. 142). Dretske says that the notion of a relevant alternative can be unpacked in terms of counterfactuals (1970, p. 144n6). RAT leads to the denial of epistemic closure principles, and that is not very attractive. One response to that is Gail Stine’s developed version of RAT. Stine’s basic idea is that the scope of relevant alternatives shifts during our familiar sceptical triad, since a sceptical possibility becomes salient and thus relevant to the epistemic subject (Stine 1976, p. 150-4). As a result of these improvements to RAT, Stine in effect developed an early version of EC<sup>k</sup> that could preserve epistemic closure yet retain the spirit of RAT (Stanley 2005, p. 20-1).

<sup>171</sup> The view has been given a slight preference by John Hawthorne (2004, chapter 4), and defended whole-heartedly by Jason Stanley (2005, chapters 5, 6, 7 and 8). Hawthorne calls it *sensitive moderate invariantism* (SMI) (2004, p. 157), while Stanley’s term is *interest-relative invariantism* (IRI) (2005, p. 85). I use Williamson’s term (*subject*) sensitive invariantism (SSI) throughout this discussion (2005a, p. 213).

<sup>172</sup> I have attached a ‘*k*’ to these abbreviations in order to remind the reader that these views are restricted to the notion of knowledge.

<sup>173</sup> See MacFarlane (2005, p. 200).

various propositions when uttered in various contexts (Bach 2005, p. 51; Williamson 2005c, p. 93). Just as speakers can employ the indexical expression ‘I’ over different contexts to refer to different persons, so knowledge-ascriptions can be used over different contexts to refer to different propositions (Stanley 2005, p. 16; Williamson 2005c, p. 93). To understand the context-sensitive notion is to master that contextual variation (Williamson 2005c, p. 93).<sup>174</sup>

When Lewis defended a version of  $EC^k$ , he thought it could explain how and why epistemology destroyed knowledge (Lewis 1996, p. 420).<sup>175</sup>  $EC^k$  can accommodate knowledge-shifts by being partly concessive to scepticism, and say that epistemology creates a particular context in which all knowledge ascriptions are false (Lewis 1996, p. 420). So on  $EC^k$  the sceptics speak truly, but only by creating special contexts of utterance in which their words express truths (Williamson 2005b, p. 688). Importantly,  $EC^k$  takes it that two speakers can consistently utter the denial as well as the affirmation of a knowledge-ascription regarding the same agent, time and situation. So the sceptic, who takes data (c)-(e)<sup>176</sup> to be valid intuitions about the nature of knowledge, can *truly* deny knowledge of S regarding  $p$  at  $t$ , while the Moorean, who takes data (a), (f) and (i) to be the valid intuitions about the nature of knowledge, can affirm knowledge of S regarding  $p$  at  $t$ . The important point from a contextualist perspective is that this might happen without the two assertors contradicting one another. As Williamson says, “They are talking past each other” (2000c, p. 94). In this sense  $EC^k$  apparently has sufficient resources to reveal the tension in the set of pre-theoretical data, since there is no longer an apparent conflict between the non-sceptical set (a), (f) and (i), and the sceptical set (c)-(e). However, their weak spot is to account for (g) and (h), since  $EC^k$  must take people’s sense of substantial disagreement over the use of context sensitive disagreements or inconsistencies to be due to some systematic error (MacFarlane 2007, p. 17). I will return to these problems in the next section.

The idea behind  $EC^k$  can be articulated accurately by using the terminology of expressing different propositions: If sentence  $P$  ascribes knowledge of propositional content  $p$  to an agent S, then if  $P$  is uttered in a non-epistemological context C, it expresses the proposition  $P^*$ , which is true if and only if  $P^*$ . However, if the sentence  $P$  is uttered in an epistemologically more sceptical context  $C^*$ , it expresses another proposition  $P^{**}$ , which is

---

<sup>174</sup> Contextualism is in this sense the anti-thesis to invariantism. That distinction is introduced and vaguely illuminated by Unger (1984, p. 248-9), as the difference between complex and simple semantics. On Hawthorne’s characterization, contextualism is the view that sentences containing the verb ‘know’ have no semantic value *simpliciter*. Rather utterances of such a sentence will have different semantic values relative to different contexts of utterance, due to contextual parameters connected to ‘know’ (Hawthorne 2004, p. 53). See previous chapter for a characterisation of Hawthorne’s use of the term ‘semantic value’. To be an *invariantist* on knowledge one must endorse that the verb ‘know’ has such a semantic value *simpliciter*. Hence an invariantist takes it that the extension of ‘know’ does *not* vary across different contexts of use (2004, p. 53, 113). My use of these terms is meant to exploit the sense that is provided by Hawthorne’s definition.

<sup>175</sup> Other prominent defenders of this view are Cohen (1998) and DeRose (1995).

<sup>176</sup> See the enumerated list of data in previous chapter.

different from  $P^*$  and which is true if and only if  $P^{**}$ . If  $P^*$  is taken to express a knowledge-relation where the epistemic standards are low (i.e. that S knows<sub>low</sub> that  $p$ ), and  $P^{**}$  is taken to express a knowledge-relation where the epistemic standards are high (i.e. that S knows<sub>high</sub> that  $p$ ), they can differ in truth-value as an ascription of S. This is evident by the fact that  $P^*$  and  $P^{**}$  are different propositions, which manifestly can differ in this respect. Hence  $P^*$  can be true when it is uttered of S regarding some time  $t$ , while  $P^{**}$  can be false. In other words, S may be correctly taken by some speaker to know<sub>low</sub> that  $p$  at  $t$ , and be taken by another speaker to not know<sub>high</sub> that  $p$  at  $t$ . Yet, the utterances are both using the same sentence  $P$ .

A comment on truth: The above commentary exploits the validity of principles (T) and (F) that are taken to elucidate the notions of truth (Williamson 1996, p. 188):

- (T)            If  $u$  says that  $P$ , then  $u$  is true if and only if  $P$ .  
(F)            If  $u$  says that  $P$ , then  $u$  is false if and only if not  $P$ .

Arguably, principles (T) and (F) are in accordance with Aristotle’s dictum, which says that, “To say of what is that it is not, or of what is not that it is, is false, while to say of what is that is, or of what is not that it is not, is true” (Williamson 1996, p. 188).<sup>177</sup> According to Williamson, this characteristic is the one Tarski sought to capture in his semantic definitions (Williamson 1996, p. 188).<sup>178</sup> Williamson does not take (T) and (F) to exhaust the notions of truth and falsity, rather he takes them to articulate what is central to our ordinary understanding of these notions (1996, p. 188). In our case an utterance using a sentence  $P$  will in context  $C$  say that  $P^*$ . Apply this to (T) and  $P^*$  is true if and only if  $P^*$ . In contrast, if uttering  $P$  in  $C^*$  says that  $P^{**}$ , then by applying (T) we get that  $P^{**}$  is true if and only if  $P^{**}$ .

We can say that on  $EC^k$  the knowledge-shift is *ascriber-dependent* since it is facts about the ascriber’s context (attention, interests, stakes, and so on) that determines which property is expressed by the ascription of knowledge to some subject (Hawthorne 2004, p. 157). One can think differently: On the sibling position  $SSI^k$ , the required shift is fixed to the *epistemic subject* in the knowledge-ascription at issue.  $SSI^k$  is also a partly concessive position vis-à-vis scepticism, where the fundamental idea is that whether or not someone knows  $p$  may be partly determined by practical facts (attention, interests, stakes, and so on)

<sup>177</sup> The quote is from Aristotle’s *Metaphysics* book  $\Gamma$  7.27.

<sup>178</sup> Tarski says that: “We should like our definition to do justice to the intuitions which adhere to the *classical Aristotelian conception of truth*” (Tarski 1944, p. 70). According to Tarski, a definition of truth must yield the following equivalence if it is to be conforming to the Aristotelian conception: “The sentence “snow is white” is true if, and only if, snow is white” (1944, p. 71). We may formalize this *Tarski definition* of what it is for a sentence  $F$  to be true in an interpretation  $M$  (i.e.  $M \models F$  where the double turnstile is to be read as “makes true”), as follows:  $M \models R(t_1, \dots, t_n)$  iff  $R^M(t_1^M, \dots, t_n^M)$ . The atomic sentence is true on the interpretation if and only if the relation we have interpreted the predicate as denoting holds of the individuals that we have interpreted the constants as denoting (Boolos, Burgess and Jeffrey 2003, p. 114).

about the *subject's* environment (Hawthorne 2004, p. 157; Stanley 2005, p. 85). It differs importantly from  $EC^k$ , since the latter is a position that takes knowledge to be context-sensitive in a strictly epistemic sense, while  $SSI^k$  allows for a sensitivity of non-epistemic factors on knowledge (Stanley 2005, p. 86).  $SSI^k$  takes the knowledge-relation to be univocal, and yet sensitive to the subject's practical situation at the time of knowing. In order to know  $p$  at  $t_1$ , the subject must rule out a different set of alternative propositions than she is required to rule out at  $t_2$  (Stanley 2005, p. 86). Consequently,  $SSI^k$  adds an additional *condition* on knowledge that involves practical facts (attention, interests, and stakes of the subject) about the subject's situation at the time of knowing (Stanley 2005, p. 86).

$SSI^k$  can explain the knowledge-shifts accordingly: When S entertains an S-hypothesis or if her life hinges on  $p$  being true, she might not know  $p$  unless she has *excessively strong evidence*. Yet, S may know that  $p$  in a context that is *epistemically similar* to these cases, but where it matters little whether  $p$  is true or not. What is important is whether the conditions on knowledge are fulfilled, and these conditions shift as the practical interests of the subject vary. Arguably, the first two cases flout some conditions on knowledge, like salience and stakes for the subject, while the latter fulfils those conditions. In this sense  $SSI^k$  can accommodate the tension between (a) and (f) versus (c)-(e), as well as make sense of the disagreement that is felt in (g). The tension is with (h), since there is no obvious reason why people should attribute knowledge only in accordance with their own present epistemic and practical circumstances. I will return to this problem in the subsequent section.

With this portrayal of  $EC^k$  and  $SSI^k$  we turn to analyse the evidence for knowledge-shifting strategies. Recall MacFarlane's abovementioned story: Williamson takes this kind of evidence to start with two intuitions. The first intuition is that in some context C it seems reasonable for a well-informed speaker to assert an affirmative knowledge-attribution of some S with respect to a true proposition  $p$ , at time  $t$  in a possible world  $w$ . Thus the well-informed speaker may utter a sentence,  $K[S, p, t, w]$ . Grant warrant<sub>ass</sub> and (1) follows (2005a, p. 214):

- (1)  $K[S, p, t, w]$  is true in C.

The second intuition is that in a different context  $C^*$  it is equally reasonable for a well-informed speaker to deny knowledge of  $S^*$  with respect to a true proposition  $p^*$  at time  $t^*$  and in world  $w^*$ . Hence our well-informed speaker may utter a sentence,  $\sim K^*[S^*, p^*, t^*, w^*]$  (2005a, p. 214). Again, grant warrant<sub>ass</sub> and (2) is true:

(2)  $\sim K^*[S^*, p^*, t^*, w^*]$  is true in  $C^*$ .<sup>179</sup>

A third feature is that both agents (S and S\*) are by traditional epistemic standards *equally epistemically well positioned* to their respective propositions ( $p$  and  $p^*$ ) in their respective time ( $t$  and  $t^*$ ) and world ( $w$  and  $w^*$ ) (Williamson 2005a, p. 215). What can differ is that S attends an epistemology seminar while S\* is not, or that matters of life and death turns on  $p$  for S while S\* is more practically unaffected by  $p^*$  (Williamson 2005a, p. 215). Arguably, these differences should not matter according to conventional epistemology, where only truth-conducive factors are relevant (Stanley 2005, p. 1). In accordance with tradition, (3) should be expected to hold (Williamson 2005a, p. 215). Speaking of tradition, knowledge should also be treated *disquotationally*, thus (4) and (5) will also hold (2005a, p. 215):

(3) If S knows  $p$  at  $t$  in  $w$ , then S\* knows  $p^*$  at  $t^*$  in  $w^*$ .

(4)  $K[S, p, t, w]$  is true in C iff S knows  $p$  at  $t$  in  $w$ .

(5)  $\sim K^*[S^*, p^*, t^*, w^*]$  is true in  $C^*$  iff S\* does not know  $p^*$  at  $t^*$  in  $w^*$ .

If we couple (1) and (2) with (4) and (5), we get (6) and (7) (2005a, p. 215):

(6) S knows  $p$  at  $t$  in  $w$ .

(7) S\* does not know  $p^*$  at  $t^*$  in  $w^*$ .

The trouble is that the last two claims constitute a clear case of inconsistency when the trio is completed by (3). For evidently, (6) is the antecedent in (3), and (7) is the negation of its consequent (Williamson 2005a, p. 215). If (6) and (7) holds, then (3) is false. And if (3) is true, then (6) or (7) (or both) must be false. Which should it be?

Let us begin with  $EC^k$ : As we have seen, the reference of the knowledge-relation can differ in C and  $C^*$  according to  $EC^k$ , so  $EC^k$  can deny (4) or (5) (or both) (2005a, p. 216). In C S knows while in  $C^*$  S\* knows\*, but to know and know\* are different relations. For instance, if knowing\* is sufficient but unnecessary<sup>180</sup> for knowing, then  $EC^k$  can take (3) as compatible with (1) and (2) at the price of denying the converse of (3) (since knowing\* requires higher epistemic standards than knowing) (2005a, p. 216). Observe that by retaining (3)  $EC^k$  remains a traditional position in the sense that the knowledge-shifts always refer to purely epistemic standards, although *which* epistemic standard it refers to can vary contextually and with

<sup>179</sup> At this point no assumption has yet been made about the structure of  $\sim K^*[S^*, p^*, t^*, w^*]$ . It might be the negation of  $K[S, p, t, w]$ , but then again it might not. The speaker too may or may not be identical (Williamson 2005a, p. 214).

<sup>180</sup> I.e. the epistemic standards are higher in  $C^*$  than they are in C.



practical interest (Stanley 2005, p. 3; Williamson 2005a, p. 216). In this sense  $EC^k$  respects *Intellectualism* by treating nothing but truth-conductive factors as relevant to whether a knowledge-relation applies (Stanley 2005, p. 6).<sup>181</sup>  $EC^k$  can respect Intellectualism and yet be flexible to accommodate the knowledge-shifts, since it can take the denotation of property by the knowledge-ascription to be dependent upon practical factors. But, once this property is denoted, the possession of it depends only on truth-conductive factors, and in this sense is knowledge not a matter of practical interests (Stanley 2005, p. 3). In our argument we see that  $EC^k$  leads to the view that the conjunction of (1), (2) and (3) entails the disjunction of negations of (4) and (5). One (or both) has to go. But to negate either one of (4) or (5) in effect entails contextualism, and, as a result, the affirmation of the conjunction (1), (2) and (3) entails  $EC^k$  (Williamson 2005a, p. 217).

On  $SSI^k$  the story is similar but with important differences:  $SSI^k$  accepts (1) and (2), and goes on to grant both (4) and (5) (Williamson 2005a, p. 217). They can do that without difficulty since the account treats knowledge as invariant. To avoid inconsistency  $SSI^k$  must deny (3) and assert both (6) and (7) (2005a, p. 217). On  $SSI^k$  the epistemic standard that S must meet in order to know is sensitive to non-epistemic features of S' circumstances (2005a, p. 217). Then  $SSI^k$  must *deny* Intellectualism and (3), and *endorse* the thesis that knowledge is also a matter of practical facts (Stanley 2005, p. 6).<sup>182</sup> Once Intellectualism is rejected,  $SSI^k$  has the resources to assert that S and S\* may be equally epistemically well positioned, yet the latter may fail to know  $p^*$  while the former knows  $p$ . Since  $p^*$  has more practical significance for S\* than  $p$  has for S,  $SSI^k$  predicts that higher epistemic standards must be met by S\* in order to know  $p^*$  than is required of S (Williamson 2005a, p. 217).

#### 6.4 Problems with the knowledge-shifting strategies

The main difference between  $EC^k$  and  $SSI^k$  is that the former exploits features in the context or situation of the *speaker* who ascribes knowledge in order to explain knowledge-shifts, while the latter exploits features in the situation of the *epistemic subject* of the knowledge ascription in order to explain the same thing (Hawthorne 2004, p. 157; Williamson 2005a, p. 217). These commitments have consequences in the sense that they predict certain types of linguistic behaviour and intuitions thereof. This section will assess their claims.

---

<sup>181</sup> Intellectualism is the orthodox position in epistemology, which takes only truth-conductive conditions to be relevant for epistemology. In this sense intellectualism is the endorsement of (3). See Stanley (2005, p. 6).

<sup>182</sup> For instance, Stanley provides a set of examples that arguably show that whether or not someone knows something may depend upon "how much is at stake" for the knowledge attributor or the putative knower (2005, p. 6).

According to Williamson, a weakness of  $SSI^k$  is that it does not assign knowledge to a well-recognized semantic category (2005a, p. 218). I think these worries can be put aside, since  $SSI^k$  is not a semantic thesis; rather it is a metaphysical thesis about the nature of the knowledge-relation (Stanley 2005, p. 120).  $SSI^k$  simply adds a condition to knowledge, and so it would be odd to hold against it that it does not attach a recognized semantical category to the relation. Arguably,  $SSI^k$  has no semantic burden to discharge and is innocent in this sense (Stanley 2005, p. 120). The matters are different on the epistemological side, since  $SSI^k$  is committed to a controversial denial of Intellectualism and endorsing a condition of practical facts on knowledge. But this is no explanatory burden on one's philosophy of language.

Another objection raised by Williamson, is that  $SSI^k$  does not classify knowing as an epistemic standard (Williamson 2005a, p. 218). Defenders of  $SSI^k$  say that two subjects may be equally "epistemically well positioned" to a proposition, yet one knows what the other fails to know. But these utterances may trigger our suspicion, since, arguably, one cannot help oneself to a firm grip of notions like justification, evidence and epistemic probability independently of knowledge.<sup>183</sup> A final initial remark is that if we identify the two propositions (1) and (2),<sup>184</sup> then  $SSI^k$  must reject either (1) or (2). By that identification (3) comes out as a tautology, which entails  $EC^k$  together with (1) and (2), and via a disjunction of the negations (4) and (5) (2005a, p. 219). That case is admittedly special, but it exploits the possibility that different competent speakers make different attributions of the same subject with respect to the same proposition, time and world (2005a, p. 219). Here  $SSI^k$  cannot mimic  $EC^k$  by establishing shift due to differences in the *subject's context*, and hence the defender of  $SSI^k$  must choose between rejecting (1) or (2). Importantly, that will attribute a false assertion on a competent speaker, and thus it constitutes an error-attribution.

The pressure to endorse (1) and (2) rests on the methodological principle of charity (MPC), where one should prefer to interpret speakers as speaking and thinking truly rather than falsely. The alleged advantage to both  $EC^k$  and  $SSI^k$  is that their ways of shifting standards attain much more flexibility for the theorist to assign knowledge with a reference that allows charitable interpretations of speakers (Williamson 2005a, p. 220). Williamson's strategy against them is to argue that both have important limits on the extent to which they are charitable, since "any reasonable view will involve the attribution to speakers of some systematic errors" (2005a, p. 220). If that is correct, *MPC* commits us to choose whatever

---

<sup>183</sup> See chapter 3. I have not treated the link between knowledge and epistemic probability.

<sup>184</sup> I.e. by letting  $S, p, t$  and  $w$  be identical with  $S^*, p^*, t^*$  and  $w^*$  respectively.

account attributes as little error as possible (Williamson 2005a, p. 220). Our objections to  $SSI^k$  are in this sense inconclusive.<sup>185</sup> Now, consider the following conversation:

ZEBRA:

Pierre: I know that this is a zebra.

Simone: How do you know that it isn't a mule cleverly painted to look like a zebra?

Pierre: For all I know it is a painted mule. So *I was wrong*. I don't know that it is a zebra after all.<sup>186</sup>

Let us seek to be maximally charitable to Pierre: What the two strategies  $EC^k$  and  $SSI^k$  can do about ZEBRA is to allocate different standards or conditions to knowing and take the first assertion to be reconciled with the last utterance of Pierre. However, the italicised passage will prove problematic for both views since that passage attributes error (Williamson 2005a, p. 220).<sup>187</sup> This is a striking fact about speakers and it confines with our data (g) that speakers “often feel caught out, corrected” (2005a, p. 221). Generally speaking, speakers in high-standard contexts tend to regard speakers in low-standard contexts as not taking possibilities of error seriously, while the latter tend to regard the former as giving too much weight to far-fetched scenarios (2005a, p. 221). Roughly speaking, “Ordinary speakers who have not consciously signed up to contextualism often have attitudes to what others say that contextualism makes inappropriate” (2005a, p. 221). ZEBRA is all the more striking since the inappropriate attitude is attributed to the speaker only separated by a slight distance of time.

Can the defenders of knowledge-shifting strategies reply? One move may be to expand the shifting strategy and say that shifts also apply to notions such as truth and falsity. This option is more natural to the defenders of  $EC^k$  since their position is part of a broader approach to philosophy of language, so I shall focus on *expansionist contextualism* (ExC). Williamson scoffs at ExC, since generalized contextualism starts to look suspiciously like “undergraduate relativism”<sup>188</sup>, which is unfortunate considering the fact that  $EC^k$  was mainly motivated in order to “tame” relativism (Williamson 2005a, p. 222; 2005c, p. 91). Another problem is that the expansionist strategy turns contextualism utterly indiscriminate, and it conforms to the philosophical adage that says, “if you reach a contradiction, draw a distinction” (Chater and Oaksford 2001, p. 203). We would therefore need independent

---

<sup>185</sup> Stanley says the same when he finds that  $EC^k$  and  $SSI^k$  fail to predict the correct intuitions in two of his cases;  $EC^k$  fails to give a correct verdict of the High Attributor-Low Subject Stakes, whereas  $SSI^k$  fails against the High Attributor-Low Subject Stakes (2005, p. 4-5, 116).

<sup>186</sup> From Williamson (2005a, p. 220).

<sup>187</sup> MacFarlane (arguing against contextualism and *pro* relativism) applies similar cases against  $EC^k$ . They are all due to what he calls “the problem of lost disagreement” since arguably  $EC^k$  evades all genuine disagreement (MacFarlane 2007, p. 18). Contextualist disagreements are shown to be on a par with the following conversation: A: “I’m tall”; B: “Hey, I’m not tall! I’m small”.

<sup>188</sup> I.e. the view that whatever a person sincerely believes is true for them.

reasons to treat truth and falsity as context-sensitive. The burden of proof is on ExC, but absent any conclusive reasons in their favour I leave this issue aside.<sup>189</sup>

Before leaving the discussion of SSI<sup>k</sup>, we need to observe that the view is in need of an extra-linguistic mechanism to explain certain phenomena. The main problem with SSI<sup>k</sup> is that it is hard to see why occurrences of “know” are necessary embedded (MacFarlane 2005, p. 202, 204). Why should people in high stakes or on epistemology seminars project *their* standard of knowledge onto others? According to Hawthorne, the *embeddedness* can be explained as the result of psychological mechanisms. Certain inferences are *anxiety provoking*, since sometimes counterpossibilities become salient during an inference from known premises, so that *in* performing the inference knowledge of the premises gets destroyed (2004, p. 161). In this sense, competently deducing a conclusion from known premises is also a way of *undermining* knowledge (Stanley 2005, p. 94). SSI<sup>k</sup> employs a *projection strategy* as a crucial step to block inferences that would otherwise create trouble for the account. So they say that sometimes apparently valid inferences provoke anxiety that renders the possibilities of mistake salient. For instance, an inference from knowledge of ordinary propositions to lottery propositions may be blocked because it is anxiety provoking and makes possibilities of error salient. The same move can block inferences made by subjects in high stakes when they assess knowledge-attributions of people in low stakes.

The conclusion thus far, is that the two views *must* attribute some degree of systematic error to competent speakers, and thus that they cannot completely satisfy the principle of charity.<sup>190</sup> As a consequence, the most charitable approach is bound to win the race. In order to locate that approach, one needs to carefully compare all sides of the explanations that postulate the error. Such a *full comparison* will not be undertaken here. Instead, I will spend the rest of this chapter discussing Williamson’s alternative response (C), and then assess whether that approach is *prima facie* preferable to EC<sup>k</sup>, SSI<sup>k</sup> and denying closure.

## 6.5 The illusion of ignorance

Williamson prefers to take a non-concessive attitude vis-à-vis the sceptic and to take knowledge to be a semantically invariant notion. In other words, he defends a version of

---

<sup>189</sup> Another suggestion is to impose restrictions on charity and only apply it to non-metalinguistic utterances (e.g. “I don’t know this is a zebra after all”) before applying it to metalinguistic assertions (e.g. “I was wrong”)? The methodology of charity tells against marginalization; speakers frequently make metalinguistic assertions. And the fact that these utterances are important tests for context-sensitivity and markers for linguistic incompetence it tells against the reply (Williamson 2005a, p. 222-3).

<sup>190</sup> Hawthorne endorses the same view when he says that: “[E]very candidate story about our puzzle has counterintuitive results” (Hawthorne 2004, p. 162). Thus we can take Hawthorne to agree with Williamson that one needs to make a considered weighting of the different options. Hawthorne is no devoted follower of SSI<sup>k</sup> though. In an interview at the University of Oslo Hawthorne explained his own attitude by saying the following: “One of these options that I explore is one that has not been explored that much. And *if you put a gun to my head* and force me to choose one of them, I would probably go for that one” (Holmen and Huvnes 2006, p. 30).

*moderate invariantism*, where he also wants to retain the validity of the closure principle. The position is called *Anti-Sceptical Insensitive Invariantism* (ASII<sup>k</sup>) (2005a, p. 225). On ASII<sup>k</sup>, the semantic value of ‘know’ is *invariant* and does *not* shift or vary across different contexts of use. It is an *anti-sceptical* position, since it takes many positive knowledge ascriptions to be true (Hawthorne 2003, p. 113). And we could add that it is *insensitive*, since whether a knowledge-relation holds is not taken to shift in accordance with the subject’s practical interests. ASII<sup>k</sup> can easily accommodate data (a) and (b), and also (f)-(i), but there is an obvious tension with the data (c)-(e). Can the resources available to ASII<sup>k</sup> explain the phenomena of scepticism and knowledge-shifts? The following section will discuss this.

We have seen that there is a close tie between epistemology and philosophy of mind on Williamson’s position. In order to explain the phenomena of scepticism and knowledge-shifts this interconnection can be brought into play, since one can attribute the shifts to familiar psychological mechanisms. As became evident above,<sup>191</sup> any mental life in the bad case is of a radically impoverished kind. How are we to understand this? The clue is Williamson’s externalism: The absence of knowledge *qua* mental state makes a difference to one’s total state of mind (2000a, p. 6). Therefore one cannot compare two cases, one bad and one good, and take the agent to have the same total mental state in both cases while knowing only in the good one (2000a p. 6, 26). A difference in knowledge constitutes a difference in mental state, but that difference is not necessarily detectable by subjects who lack knowledge (2000a, p. 26). Because our minds are characteristically anti-luminous, we *can* be in different mental states in different cases between which we cannot discriminate (2000a, p. 26).

This is not in itself a problem for the sceptic, but we can utilize the point: The sceptic exploits this feature of our minds by manipulating psychological mechanisms as well as features in the dialectical situation. In a discourse one can challenge any belief *p* with the effect that one temporarily suspend that belief in order to recover it (2000a, p. 2).<sup>192</sup> This is the *suspension test*, which is associated with a discursive mode of reasoning that seeks to convince an opponent who does not already accept the belief in question. In order to convince one’s opponent without begging the question, one needs a dialectical starting-point accepted by the opponent (2000a, p. 2).<sup>193</sup> The sceptic makes use of that feature when she calls knowledge of *p* into question. As a result, one provisionally treats *p* as though it did not belong to our body of knowledge in order to support it on independent ground (2000a, p. 10).

---

<sup>191</sup> See section 6.1.

<sup>192</sup> E.g. by supporting the proposition with an appropriate argument to the effect that *p* is true.

<sup>193</sup> Plato endorses this feature of discourse when he let Socrates be subjugated to similar rules in his dialogues. As Socrates puts it: “If you remember, when I was answering you about shape, we rejected the kind of answer that tried to answer in terms *still being the subject of inquiry* and not yet agreed upon” (Plato 2005, p. 211).

One assesses  $p$  relative to the rest of one's knowledge (or evidence), and if the rest of one's knowledge (or evidence) justifies it  $p$  is recovered (2000a, p. 10). Importantly, a temporary suspension of  $p$  is consistent with retaining one's belief. In this sense, truth is differentiated from discursive conditions (Lewis 1984, p. 63; Williamson 2000a, p. 2).

The suspension test yields poor results when too much knowledge is simultaneously called into question, so this is exactly what the sceptic does (2000a, p. 10). On a bundle of sceptical strategies one exploits the dialectical effects of challenging propositions. Since one is not entitled to rely on what is challenged, the sceptic can effectively impoverish the discourse by repeating and expanding her challenges (Williamson 2000a, p. 188). As Lewis says, "The rules of disputation sometimes give the wrong side a winning strategy. In particular, they favour the sceptic" (1984, p. 63). The sceptic relies uncritically on rules of dialectical engagement that have evolved in order to serve other practical purposes. By so doing the sceptic employs rules completely out of proportion to the practical tasks they were developed for, and there is no guarantee that this will track the truth (2000a, p. 188).

ASII<sup>k</sup> must be able to explain the abovementioned data in a charitable manner competing with SSI<sup>k</sup> and EC<sup>k</sup>. Williamson's strategy is to explain the problematic evidence as effects stemming from *psychological bias effects* caused by differences in salience (2005a, abstract). Recall the above analysis: If we assume that S knows  $p$  at  $t$  in  $w$ , then on ASII<sup>k</sup> (1) and (3)-(6) are true while (2) and (7) is rejected. The assertions of knowledge in low standards are (typically) true while the denials of knowledge in high standards are false (2005a, p. 225). What explains the persuasiveness of scepticism and the illusion of ignorance is psychology. The bearing thought is that appearance of ignorance undermines knowledge in a way in which the appearance of knowledge does not undermine ignorance (2000a, p. 265). Whether S knows  $p$  depends on a multitude of considerations, which rarely point in the same direction and that are vulnerable to psychological biases (2005a, p. 225; 2005b, p. 695). Whether one should believe  $p$  typically depends on probability and similarity to cases of false belief, and with a finite mind we have no choice but to judge on impressionistic basis whether one knows that  $p$  (2005a, p. 225).<sup>194</sup> <sup>195</sup> When we make judgements about these matters, we apply convenient but highly fallible *rules of thumb* rather than algorithms. Those rules make us liable to bewilderments by psychological salience. The sceptical cases are construed in such a way that they turn the focus of the ascriber's attention on considerations that go *against*

---

<sup>194</sup> This is related to the idea that knowledge is governing belief. In order to know  $p$ , it is required that the true belief that  $p$  have margins for error. See chapter 3.

<sup>195</sup> As Williamson says, "Suppose that we do apply the concept of knowledge on the basis of something more like a fallible recognitional capacity than a conceptual analysis" (2005b, p. 697).

knowledge-ascriptions. Sceptics achieve these effects by providing a sceptical hypothesis, a vivid story or by raising practical stakes (2005a, p. 225-6).

It is important to grasp the dialectics here: The fundamental point is that *if* one can establish that a set of intuitions  $\Gamma$  is the result of salience effects (that elicit inferences of a kind that are highly dependent upon idiosyncratic features of descriptions of the background situation), *then*  $\Gamma$  is not the kind of evidence that shed light on the nature of knowledge (Stanley 2005, p. 9). Cases of framing effects are special, since they do not follow discernible patterns and fail to reflect plausible general claims about knowledge and rationality. Strictly speaking, they are unexpected results usually interpreted as some kind of irrationality (Stanley 2005, p. 9).<sup>196</sup> To categorize the sceptical intuitions or data (c)-(e) as resulting from bias effects is to partake in a *projectivist strategy* that explains mistaken intuitions about knowledge by appeal to some psychological feature of speakers (Stanley 2005, p. 99).

We can safely conclude that ASII<sup>k</sup> is an anti-sceptical *error theory* that seeks to explain our systematic errors (c)-(e) through psychology. The question is whether that psychological explanation is convincing. In the next section I will say something about to what extent cognitive science and psychology can be expected to confirm Williamson's claim.

### **6.6 The possibility of empirical support for the illusion of ignorance**

Any serious appeal to psychological mechanisms, like salience- and framing effects, needs support from scientific evidence. Williamson makes no attempt to support his own speculations empirically, but lingers in the philosophical armchair. Perhaps there is no harm in that, but in order to decide whether Williamson's story explains speaker-error we need to confirm empirically those propositions.

There have been generated much evidence for the effects of framing- and salience effects on people's moral considerations, but only recently have researchers oriented themselves towards epistemological considerations. Hence are the results yet to come.<sup>197</sup> There is still some reason to think that Williamson's claims can be confirmed, because researchers have found analogous results on *responsibility-attributions*. These experiments suggest a possible set-up for a test on knowledge-attributions.

Lerner, Goldberg and Tetlock thought that justice-judgments could be influenced by emotive conditions in a subject. Their hypothesis was that anger elicits simple heuristic modes

---

<sup>196</sup> See Kahneman and Tversky (1982), and Kahneman (2003, p. 1458-60) for the interpretation that framing effects reveal irrationality. This interpretation is contested by Chater and Oaksford when they take the results to indicate that everyday rationality is probabilistically based, and hence can accommodate better the effects of uncertainty revealed by the framing effect cases (2001, p. 204)

<sup>197</sup> I owe this observation to Shaun Nichols, who is far more familiar with the empirical findings than the present author.

of information processing that will influence people's responsibility-attributions in predictive ways (1998, p. 563, 570). Importantly, the emotive state of the subject was predicted to influence the subject unbeknownst to her. Metaphorically speaking, the emotive condition is like a "mental contamination" in which implicit desires to punish operate outside of conscious awareness causing people to use irrelevant information in their judgements (1998, p. 564). *Qua* mental contamination it was further predicted by the scientists that subjects might be triggered to apply more systematic and self-critical modes of thinking that were less prone to the impact of irrelevant information and situational predictors. Thus some were informed that "experts" should assess them. By hypothesis this would elicit subjects' accountability for their verdicts and trigger a self-critical mode (1998, p. 564, 570).

The result was that anger-primed participants made more punitive attributions than the non-conditioned participants thus confirming that anger influenced the participants (1998, p. 568, 570). The participants were ignorant about the effects of the anger-eliciting situation and they also agreed with the set-up that the judgements were totally unrelated to the anger-eliciting situation. In this sense, they agreed that the anger was a piece of mental contamination. On the positive side they found that by eliciting people's accountability one could attenuate punitiveness. This was an effect even though the participants still reported being somewhat angry through the whole experiment due to the anger-eliciting situation. In other words, the results indicate that just as emotion can contaminate thought and reasoning, so can styles of thought *immunize* the subject. Interestingly, this effect remained over different tasks and contexts, and one can speak of an "accountability carryover" (1998, p. 569, 570-2).

What are important for us are the prospects of similar effects, both contaminating and immunizing, on knowledge-attributions. This kind of experiment has not been conducted on knowledge, but one could do it by simply modelling the set-up on Lerner, Goldberg and Tetlock (1998, p. 569). One would need to locate an ignorance-eliciting mental state that would prompt the application of simple heuristic modes vulnerable to irrelevant information made salient by the ignorance-eliciting state. Plausibly the mental condition is not required to be affective.<sup>198</sup> Then our hypothesis would be that this state would bring about a simplification of thought that makes the subjects more prone to attribute ignorance rather than knowledge. We would also need to elicit the effect without the participants relating the emotional-eliciting state to the knowledge-attributions, which would make it possible to elicit accountability in order to motivate more self-reflectiveness. Then the subjects could be expected to moderate the effects of non-affective frames on their thinking.

---

<sup>198</sup> I owe this point to Shaun Nichols.



In order to establish ignorance as the result of psychological biases, we need to parallel the results on responsibility-attributions. The challenge is to find the corresponding non-affective frame that, just like anger, can influence knowledge-attributions. Since the jury is not even set on this trial Williamson's story cannot be flat-out asserted.

### 6.7 Knowledge and practical reasoning

We have seen that dialectical rules can be a poor guide to epistemology (Williamson 2000a, p. 22). The question is: Are there any reliable guides? According to several authors, an intuitive and attractive link between knowledge and practical reasoning (or action) may prove more useful to a theory of knowledge than dialectical practice (Hawthorne 2004, p. 29; Stanley 2005, p. 9, 12; Williamson 2000a, p. 15). That is captured in the following principle:

(KPR) A first-person present-tense ascription of 'know' with respect to a proposition is true in a context iff that proposition is an *appropriate premise* for practical reasoning in that context.<sup>199</sup>

The idea is to employ *KPR* as a constraint on knowledge-ascriptions. Note that this might be used to reinforce the arguments for shifting standards (2005a, p. 227). Suppose *S* and *S\** deliberate in order to commence some action  $\phi$ ; then by *KPR* we know both (13) and (14). Assume also that *p* makes little practical difference to the speaker in *C*, while *p\** makes an enormous practical difference to the speaker in *C\**; then by *KPR* we get (15) and (16):

(13)  $K[S, p, t, w]$  is true in *C* iff *p* is an appropriate premise for practical reasoning in *C*.

(14)  $\sim K^*[S^*, p^*, t^*, w^*]$  is true in *C\** iff *p\** is not an appropriate premise for practical reasoning in *C\**.

(15) *p* is an appropriate premise for practical reasoning in *C*.

(16) *p\** is not an appropriate premise for practical reasoning in *C\**.

That reinforces the knowledge-shifting strategies (*KSS*), because  $ASII^k$  must reject at least one of *KPR*, (15) or (16). Which shall it be? Observe that without *KPR* knowledge "would lose some of its significance", which runs counter to data, namely (b) (2005a, p. 228). It is also a very intuitive thought that what is appropriate to use as a premise in practical reasoning at least partly depends on how much is at stake (Stanley 2005, p. 6; Williamson 2005a, p. 228-9). How can  $ASII^k$  cope with *KPR* and the tension with (15) and (16)?

Before we answer that question: Observe that neither  $EC^k$  nor  $SSI^k$  can accommodate *KPR* without friction (Williamson 2005a, p. 229). For suppose that you attend a seminar on

---

<sup>199</sup> See Williamson (2005a, p. 15).

radical scepticism, where due to the alleged shifts one cannot truly say, “I have hands”. Suppose that in the middle of the seminar someone offers you a bet on which you gain 5£ if you have hands and lose 50£ otherwise. According to both  $EC^k$  and  $SSI^k$ , you should not have an appropriate premise to use in a sensible practical reasoning where you rationally accept the bet (Williamson 2005a, p. 229). But to accept such a bet *is* very sensible, and accordingly the proposition that you have hands is an appropriate premise. Then that proposition *must* be true, unless this case is counterexample to  $KPR$ . That is yet another reason to reject  $EC^k$  and  $SSI^k$ .

We must still explain the shift on  $ASII^k$ . Recall that we established anti-luminosity for non-trivial conditions in chapter 2: Since whether  $p$  is an appropriate premise for practical reasoning *is* highly non-trivial that too is a non-luminous condition. It is one thing that  $p$  has the required authority and is an appropriate premise for S; another question is whether S *knows* that  $p$  has this property (2005a, p. 230). Williamson’s suggestion is that in some cases where  $p$  is appropriate but S does *not* know it, we may still criticize S depending on how much is at stake (2005a, p. 230). This is captured in the following reformulation of  $KPR$ :

(KPR+) One knows  $p$  iff  $p$  is an appropriate premise for one’s practical reasoning.<sup>200</sup>

Just like S could know  $p$  without knowing that she knows (chapter 3), S may have authority to use  $p$  as a premise in one’s practical reasoning *without knowing* that she has (2005a, p. 231). Importantly, we can apply  $KPR+$  to explain why both (15) and (16) seem so plausible: Suppose that S in C knows that she knows  $p$ , while  $S^*$  in  $C^*$  is ignorant about whether he knows  $p^*$ . Suppose further that in  $C^*$  stakes are higher than in C. It seems plausible to judge more sternly the failure of second-order knowledge in  $C^*$  than in C, and in that sense  $p$  is appropriate for S in C while  $p^*$  is inappropriate for  $S^*$  in  $C^*$  (2005a, p. 231). More generally: How many orders of knowledge are required depends on how much is at stake (2005a, p. 231). Observe that *if* this move is acceptable it provides  $ASII^k$  with a systematic response to arguments from practical differences to shifting semantic standards.<sup>201</sup>

Let me illustrate: Both  $EC^k$  and  $SSI^k$  takes it that to utter, “I know that the plane stops in Chicago”, after a causal glance at the timetable is true when uttered by Lo in  $C_{low}$ , but false as uttered by Hi in  $C_{high}$ . The pressure is on  $ASII^k$  to deny one of these utterances as false and lose a corresponding amount of charity. According to Williamson, the utterances of Lo and Hi are close to the boundary between knowledge and ignorance, and then anti-luminosity

<sup>200</sup> See Williamson (2005a, p. 231).

<sup>201</sup> These “orders of knowledge” are the result of iterations of knowledge. As an effect of the iteration one requires another margin for error, which makes the requirements on belief stricter. See chapter 3.

predicts that one may know without being in a position to know that one knows (2005a, p. 232). So it is plausible that both Lo and Hi know that the plane stops in Chicago, but that neither knows that he/she knows. Since stakes are higher for Hi than Lo this lack of second-order knowledge is judged more sternly for Hi. Given *KPR+* the premise is appropriate for both subjects, but Hi has more reason than Lo to check out his practical reasoning, and hence ensure knowledge of his knowledge. That requires Hi to engage in second-order practical reasoning about whether to trust the first-order reasoning (2005a, p. 232). Since Hi does not know that the first-order premise is appropriate, the second-order premise that the first order premise is appropriate is inappropriate (though true) (2005a, p. 233). Because Hi is required to ensure this and Lo is not, we judge Hi more sternly than Lo for the lack of second-order appropriateness. *That move* explains the shift and why both (14) and (15) are so plausible.

What happens when Hi and we assess Hi's situation and find second-order knowledge wanting is that we are *struck* by the disastrous consequences of believing falsely. As an effect, the weakness of Hi's (and Lo's) epistemic position becomes salient, and we (and Hi) tend to retract the knowledge-attribution of both Hi and Lo. Again, to explain this Williamson invokes the abovementioned psychological story. Psychological bias effects tend to push us all in the direction of *overkill* by denying knowledge of Hi and Lo (2005a, p. 234-5).

## 6.8 Conclusion

It is time to wrap up our discussion: We found that denying closure was almost as unattractive as embracing scepticism, and so we decided to consider some other options. Both  $EC^k$  and  $SSI^k$  looked promising, but we found problems with them:  $EC^k$  had to explain competent speaker's attribution of error (*retraction*), and the feeling of *disagreement* as systematic errors, while  $SSI^k$  had to attribute error as well as employ salience effects to explain why occurrences of know were *embedded* in the here-and-now.

We considered Williamson's position  $ASII^k$ , and found that the challenge for  $ASII^k$  is the apparent variability of standards suggested by the knowledge-shifting arguments. To explain those shifts a story is required, so Williamson appeals to psychological salience- and bias effects. By coupling those mechanisms with plausible requirements on practical reasoning and his claims about anti-luminosity and margins for error he was able to explain why those shifts were errors. Compared with Dretske Williamson provides a more attractive position, since he can retain *intuitive* closure while avoiding scepticism. A comparison with the other views shows that  $ASII^k$  can accommodate the data that was disturbing to  $EC^k$ , namely (g), and  $SSI^k$ , namely (h). The price is that the data about sceptical knowledge-shifts

(c) becomes problematic. Williamson explains (c) as an error of attribution resulting from sceptical dialectics and salience effects. The apparent knowledge-shifts that result from sceptical influence and high stakes are due to requirements of higher-order knowledge. These requirements may *seem* to invalidate first-order knowledge, but those appearances stem from salience effects. In general, does higher-order ignorance not warrant the rejection of first-order knowledge (Williamson 2005a, p. 234). Of course, all views need to invoke some degree of systematic error. The error invoked on  $EC^k$  lacks a plausible explanation to back it up, while  $SSI^k$  invokes salience effects in a manner analogous to  $ASII^k$ . But, whereas that move seems *ad hoc* and out of place on  $SSI^k$ , it is an integrated part of Williamson's account. The psychological mechanisms are a natural consequence of taking knowledge to be a purely mental state coupled with a conception of mind that takes mental conditions to be constitutively dependent on external features and characteristically anti-luminous. We also saw that neither  $EC^k$  nor  $SSI^k$  could accommodate the plausible connection between knowledge and practical reasoning (i.e. the principle of *KPR*) without friction, but that  $ASII^k$  had the explanatory resources available to explain the appearances of knowledge-shifts in a systematic manner. An important role for knowledge is to aid our practical reasoning, so any account of knowledge should be able to accommodate something like *KPR*.

So compared with  $EC^k$  and  $SSI^k$  Williamson's position  $ASII^k$  is preferable. Admittedly, the empirical results to confirm that psychological explanation, are still wanting, so future evidence may turn against  $ASII^k$ . Yet, the current evidence support that we take  $ASII^k$  to be our preferred view. All in all, I take that to provide strong support for Williamson's position  $ASII^k$ . Ignorance is a natural human state, and its presence tends to outstrip the appearance of knowledge. In unfavourable circumstances that creates an illusion of ignorance, and we must make sure that we do not mistake those illusions for the real thing. Knowledge *is* real.

## Bibliography:

- Aristotle 2005a, *Metaphysics*, in Cohen, S.M., P. Curd, and S.D.C. Reeve (eds.) 2005, *Readings in Ancient Greek Philosophy: From Thales to Aristotle*, 3<sup>rd</sup> ed., Hackett Publishing Company, Indianapolis, p. 758-808.
- 2005b, *Posterior Analytics*, in Cohen, S.M., P. Curd, and S.D.C. Reeve (eds.) 2005, *Readings in Ancient Greek Philosophy: From Thales to Aristotle*, 3<sup>rd</sup> ed., Hackett Publishing Company, Indianapolis, p. 676-93.
- Austin, J.L. 1963, *How to Do Things with Words*, Oxford University Press, Oxford.
- 1979, 'Performative Utterances', in Urmson, J.O. and G.J. Warnock (eds.), *Philosophical Papers*, 3. ed., Oxford University Press, Oxford, p. 233-52, in Martinich, A.P. (ed.) 2001, *The Philosophy of Language*, Oxford University Press, Oxford, p. 130-9.
- Bach, K. 2005, 'The Emperor's New 'Knows'', in Peter, G. and G. Preyer (eds.) 2005, *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, Oxford University Press, Oxford, p. 51-89.
- Beaney, M. 2003, 'Analysis', *Stanford Encyclopedia*, <<http://plato.stanford.edu/entries/analysis/>>.
- Boolos, G.S., J.P. Burgess and R.C. Jeffrey 2003, *Computability and Logic*, 4<sup>th</sup> ed., Cambridge University Press, Cambridge.
- Brandom, R. 1994, *Making It Explicit*, Harvard University Press, Cambridge.
- 1995, 'Knowledge and the Social Articulation of the Space of Reasons', in *Philosophy and Phenomenological Research* 55, p. 895-908, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 424-32.
- 2000, *Articulating Reasons: An Introduction to Inferentialism*, Harvard University Press, Cambridge.
- 2002, *Tales of the Mighty Dead*, Harvard University Press, Cambridge.
- Burge, T. 1979, 'Individualism and the Mental', in French, P., T. Uehling and H. Wettstein (eds.) 1979, *Studies in Metaphysics*, University of Minnesota Press, Minnesota, in Chalmers, D.J. 2002, *Philosophy of Mind: Classical and Contemporary Readings*, Oxford University Press, Oxford, p. 597-607.
- Bruckner, A. 1992, 'Semantic Answers to Skepticism', in *Pacific Philosophical Quarterly* 73, University of Southern California, p. 200-19, in DeRose, K. and T.A. Warfield (eds.) 1999, *Skepticism: A Contemporary Reader*, Oxford University Press, Oxford, p. 131-44.
- Cappelen, H. 2005, 'Pluralistic Skepticism', in Hawthorne, J. (ed.) 2005, *Philosophical Perspectives: Epistemology*, vol. 19, Blackwell Publishing, Oxford, p. 15-39.
- Cartwright, N. 1980, 'Do the Laws of Physics State the Facts?', in *Pacific Philosophical Quarterly* 61, 1980, p. 75-84, in Curd, M. and J.A. Cover (eds.) 1998, *Philosophy of Science: The Central Issues*, W.W. Norton & Company, New York, p. 865-77.
- Chater, N. and M. Oaksford 2001, 'Human Rationality and the Psychology of Reasoning: Where do we go from here?', in *British Journal of Psychology*, vol. 92, The British Psychological Society, London, p. 193-216.
- Cohen, S. 1998, 'Contextualist Solutions to Epistemological Problems: Scepticism, Gettier, and the Lottery', in *Australasian Journal of Philosophy* 76, 2, 1998, p. 289-306, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 517-30.
- Cohen, J. and P.D. Magnus 2003, 'Williamson on Knowledge and Psychological Explanation', *Philosophical Studies* 116 (1), p. 37-52.
- Dancy, J. 1985, *Introduction to Contemporary Epistemology*, Blackwell Publishing, Oxford.
- Davidson, D. 1968, 'On Saying That', in Davidson, D. 2001, *Inquiries into Truth and Interpretation*, 2<sup>nd</sup> ed., Oxford University Press, Oxford, p. 93-108.
- 1970, 'Mental Events', in Davidson, D. 2001, *Essays on Actions and Events*, Oxford University Press, Oxford, p. 207-27.
- 1973, 'Radical Interpretation', in Davidson, D. 2001, *Inquiries into Truth and Interpretation*, 2<sup>nd</sup> ed., Oxford University Press, Oxford, p. 125-39.
- 1974, 'On the Very Idea of a Conceptual Scheme', in Davidson, D. 2001, *Inquiries into Truth and Interpretation*, 2<sup>nd</sup> ed., Oxford University Press, Oxford, p. 183-98.
- 1983, 'A Coherence Theory of Truth and Knowledge', in Davidson, D. 2001, *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford, p. 137-53.
- 1987, 'Knowing One's Own Mind', in Davidson, D. *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford, p. 15-38.
- DeRose, K. 1995, 'Solving the Skeptical Problem', in *The Philosophical Review* 104, 1, 1995, p. 1-7, 17-52, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 482-502.
- 1999, 'Introduction: Responding to Skepticism', in DeRose, K. and T.A. Warfield (eds.) 1999, *Skepticism: A Contemporary Reader*, Oxford University Press, Oxford, p. 1-24.

- 2002, 'Review: Timothy Williamson: Knowledge and its Limits', in *British Journal of Philosophy of Science*, 53, 2002, British Society for the Philosophy of Science, p. 573-77.
- Descartes, R. 1998, *Meditations on First Philosophy*, in Ariew, R. and E. Watkins (eds.) 1998, *Modern Philosophy: An Anthology of Primary Sources*, Hackett Publishing Company, Cambridge, p. 22-55.
- Douven, I. and T. Williamson 2006, 'Generalizing the Lottery Paradox', in *The British Journal of Philosophy and Science*, 57, Oxford University Press, Oxford, p. 755-79.
- Dretske, F. 1970, 'Epistemic Operators', *The Journal of Philosophy* 67, p. 1007-23, in DeRose, K. and T.A. Warfield (eds.) 1999, *Scepticism: A Contemporary Reader*, Oxford University Press, Oxford, p. 131-44.
- 1985, 'Précis of Knowledge and the Flow of Information', in Kornblith, H. (ed.), *Naturalizing Epistemology*, MIT Press, Cambridge.
- 2005, 'Is Knowledge Closed under Known Entailment? The Case Against Closure' and 'Reply to Hawthorne', in Steup, M. and E. Sosa (eds.), *Contemporary Debates in Epistemology*, Blackwell Publishing, Oxford, p. 13-26 and 43-6.
- Dummett, M. 1981, *Frege: Philosophy of Language*, 2<sup>nd</sup> ed., Gerald Duckworth & Company Limited, London.
- 1993, 'Mood, Force, and Convention', in Dummett, M. 1993, *The Seas of Language*, Oxford University Press, Oxford, p. 202-23.
- Engel, P. 2002, 'What Can We Learn from Psychology about the Nature of Knowledge?', *Rivista di filosofia neo-scholastica*, XCIV, 4, p. 725-40.
- Forbes, G. 1995, 'Realism and Skepticism: Brains in a Vat Revisited', in *Journal of Philosophy* 92, p. 205-22, in DeRose, K. and T.A. Warfield (eds.) 1999, *Scepticism: A Contemporary Reader*, Oxford University Press, Oxford, p. 61-75.
- Frege, G. 1879, *Begriffsschrift*, in Beaney, M. (ed.) 1997, *The Frege Reader*, Blackwell Publishing, Oxford, p. 47-78.
- 1892, 'On Sinn and Bedeutung', in *Zeitschrift für Philosophie und philosophische Kritik*, 100, p. 25-50, in Beaney, M. (ed.) 1997, *The Frege Reader*, Blackwell Publishing, Oxford, p. 151-71.
- 1918, 'Thought', in *Beiträge zur Philosophie des deutschen Idealismus I*, p. 58-77, in Beaney, M. (ed.) 1997, *The Frege Reader*, Blackwell Publishing, Oxford, p. 325-45.
- 1949, 'On Sense and Nominatum', in Feigl, H. and W. Sellars (eds.) 1949, *Readings in Philosophical Analysis*, Appleton-Century-Crofts, New York, p. 85-102, in Martinich, A.P. 2001, *The Philosophy of Language*, 4<sup>th</sup> ed., Oxford University Press, Oxford, p. 199-211.
- 1997, 'Logic', in Beaney, M. (ed.) 1997, *The Frege Reader*, Blackwell Publishing, Oxford, p. 227-50.
- Gettier, E.L. 1963, 'Is Justified True Belief Knowledge?', in *Analysis*, 1963, p. 121-3, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 58-9.
- Goldie, P. 2000, *The Emotions: A Philosophical Exploration*, Oxford University Press, Oxford.
- Goldman, A. 1976, 'Discrimination and Perceptual Knowledge', in *The Journal of Philosophy*, Vol. 73, No. 20, p. 771-91.
- Grice, P. 1989, *Studies in the Way of Words*, Harvard University Press, Cambridge.
- Harman, G. 2002, 'Reflections on Knowledge and Its Limits', in *The Philosophical Review*, Vol. 111, No. 3., p. 417-28.
- Hawthorne, J. 2004, *Knowledge and Lotteries*, Oxford University Press, Oxford.
- 2005, 'Is Knowledge Closed under Known Entailment? The Case for Closure', in Steup, M. and E. Sosa (eds.), *Contemporary Debates in Epistemology*, Blackwell Publishing, Oxford, p. 26-43.
- Hintikka, J. 1962, *Knowledge and Belief: An Introduction to the Logic of the two Notions*, Cornell University Press, London, reprint edition 2005, King's College Publications, London.
- Holmen, H.A. and T.T. Huvenes 2006, 'Doing Metaphysics as if Kant Never Happened: An Interview with John Hawthorne', in *Filosofisk supplement* vol. 2, 2006, Bookline Trykkeri, Drøbak, p. 27-31.
- Hume, D. 1969, *A Treatise of Human Nature*, Penguin Books, London.
- 1993, *An Enquiry Concerning Human Understanding*, 2<sup>nd</sup> ed., Hackett Publishing Company, Indianapolis.
- Kahane, H. and P. Tidman 2003, *Logic and Philosophy: A Modern Introduction*, 9<sup>th</sup> ed., Thomson/Wadsworth, Belmont.
- Kahneman, D. 2003, 'Maps of Bounded Rationality: Psychology for Behavioral Economics', in *American Economic Review* vol. 93 no. 5, American Economic Association, p. 1449-75.
- Kahneman, D. and A. Tversky 1982, 'Variants of Uncertainty', in *Cognition*, Elsevier Sequia S.A., Lausanne, 11, p. 143-57.
- Kaplan, D. 1978, 'Dthat', in Cole, P. (ed.) 1978, *Syntax and Semantics*, vol. 9, Academic Press, New York, p. 221-53, in Martinich, A.P. (ed.) 2001, *The Philosophy of Language*, 4. ed., Oxford University Press, Oxford, p. 325-38.
- Kim, J. 1998, *Philosophy of Mind*, Westview Press, Oxford.
- King, J. 1998, 'What is a Philosophical Analysis?', in *Philosophical Studies* vol. 90, nr 2, Springer Netherlands, p. 155-79.

- Klein, P. 1971, 'A Proposed Definition of Propositional Knowledge', *The Journal of Philosophy* 68, 16, p. 471-82, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 60-6.
- Kripke, S. 1979, 'A Puzzle About Belief', in Margalit, A. (ed.) 1979, *Meaning and Use*, D. Reidel, Dordrecht, p. 239-83, in Martinich, A.P. (ed.) 2001, *The Philosophy of Language*, 4<sup>th</sup> ed., Oxford University Press, Oxford, p. 405-34.
- 1981, *Naming and Necessity*, Blackwell Publishing Ltd, Oxford.
- Lerner, J.S., J.H. Goldberg and P.E. Tetlock 1998, 'Sober Enough: The Effects of Accountability, Anger, and Authoritarianism on Attributions of Responsibility', *Personality and Social Psychology Bulletin* 1998, vol. 24, Sage Publications, London, p. 563-74.
- Lewis, D. 1968, 'Counterpart Theory and Quantified Modal Logic', *Journal of Philosophy* 65, p. 113-26.
- 1984, 'Putnam's Paradox', in *The Australasian Journal of Philosophy* 62, 1984, p. 221-36, in Lewis, D. 1999, *Papers in Metaphysics and Epistemology*, Cambridge University Press, Cambridge, p. 56-77.
- 1996, 'Elusive Knowledge', in *The Australasian Journal of Philosophy* 74, 1996, p. 549-67, in Lewis, D. 1999, *Papers in Metaphysics and Epistemology*, Cambridge University Press, Cambridge, p. 418-45.
- Locke, J. 1998, *Essay Concerning Human Understanding*, in *Works*, 1823, London, in Ariew, R. and E. Watkins (eds.) 1998, *Modern Philosophy: An Anthology of Primary Sources*, Hackett Publishing Company, Cambridge, p. 270-373.
- Longino, H. 2002, *The Fate of Knowledge*, Princeton University Press, Princeton.
- Ludlow, P. 2005, 'Contextualism and the New Linguistic Turn in Epistemology', in Peter, G. and G. Preyer (eds.) 2005, *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, Oxford University Press, Oxford, p. 11-50.
- MacFarlane, J. 2005, 'The Assessment Sensitivity of Knowledge Attributions', in Gendler, T.S., J. Hawthorne (eds.) 2005, *Oxford Studies in Epistemology* vol. 1, Oxford University Press, Oxford, p. 196-233.
- 2007, 'Relativism and Disagreement', in *Philosophical Studies* 132, Springer Netherlands, p. 17-31.
- McCulloch, G. 1995, *The Mind and its World*, Routledge, London.
- Moore, G.E. 1959, 'Certainty', in Moore, G.E. 1959, *Philosophical Papers*, George Allen & Unwin Ltd, London, p. 227-51.
- 1962a, *Commonplace Books*, George Allen & Unwin Ltd, London.
- 1962b, 'Four Forms of Scepticism', in Moore, G.E. 1962, *Philosophical Papers*, Collier Books, New York, p. 220-2, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 27-8.
- 1962c, 'Proof of an External World', in Moore, G.E. 1962, *Philosophical Papers*, Collier Books, New York, p. 144-8, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 24-6.
- 1966, *Ethics*, 2<sup>nd</sup> ed., Oxford University Press, London.
- Nozick, R. 1981, 'Knowledge and Skepticism', in Nozick, R. 1981, *Philosophical Explanations*, Harvard University Press, Cambridge, p. 172-85, 197-217, in Kim, J. and E. Sosa (eds.) 2000, *Epistemology: An Anthology*, Blackwell Publishing, Oxford, p. 79-101.
- Pagin, P. 2007, 'Assertion', Stanford Encyclopedia of Philosophy, < <http://plato.stanford.edu/entries/assertion/> >.
- forthcoming, 'Against the Knowledge Account of Assertion'.
- Parmenides 2005, 'Parmenides', in Cohen, S.M., P. Curd, and S.D.C. Reeve (eds.) 2005, *Readings in Ancient Greek Philosophy: From Thales to Aristotle*, 3<sup>rd</sup> ed., Hackett Publishing Company, Indianapolis, p. 35-41.
- Perry, J. 2001, *Reference and Reflexivity*, CSLI Publications, Stanford.
- Peter, G. and G. Preyer 2005, 'Introduction: The Limits of Contextualism', in Peter, G. and G. Preyer (eds.) 2005, *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, Oxford University Press, Oxford, p. 1-7.
- Plato 2005, *Meno*, in Cohen, S.M., P. Curd, and S.D.C. Reeve (eds.) 2005, *Readings in Ancient Greek Philosophy: From Thales to Aristotle*, 3<sup>rd</sup> ed., Hackett Publishing Company, Indianapolis, p. 203-28.
- Pryor, J. 2005, 'Is There Immediate Justification?', in Steup, M. and E. Sosa (eds.), *Contemporary Debates in Epistemology*, Blackwell Publishing, Oxford, p. 181-202.
- Putnam, H. 1973, 'Meaning and Reference', *Journal of Philosophy* 70, p. 699-711, in Martinich, A.P. (ed.) 2001, *The Philosophy of Language*, 4<sup>th</sup> ed., Oxford University Press, Oxford, p. 288-95.
- 1975, 'The Meaning of "Meaning"', in Gunderson, K. (ed.) 1975, *Language, Mind and Knowledge*, University of Minnesota Press, Minnesota, p. 131-93, in Chalmers, D.J. 2002, *Philosophy of Mind: Classical and Contemporary Readings*, Oxford University Press, Oxford, p. 581-96.
- 1981, 'Brains in a Vat', in *Reason, Truth and History*, Cambridge University Press, Cambridge, in DeRose, K. and T.A. Warfield (eds.) 1999, *Scepticism: A Contemporary Reader*, Oxford University Press, Oxford, p. 27-42.

- Quine, W.V.O. 1960, *Word and Object*, The M.I.T. Press, Cambridge.
- Radford, C. 1966, 'Knowledge – by Examples', in *Analysis*, 27, p. 1-11.
- 1988, 'Radford Revisiting', in *The Philosophical Quarterly*, Vol. 38, No. 153, p. 496-99.
- Reed, B. 2006, 'Shelter for the Cognitively Homeless', in *Synthese* vol. 148 nr. 2 2006, Springer Publishing, Netherlands, p. 303-8.
- Russell, B. 1905, 'On Denoting', *Mind* 14, p. 479-93, in Martinich, A.P. (ed.) 2001, *The Philosophy of Language*, 4<sup>th</sup> ed., Oxford University Press, Oxford, p. 212-20.
- 1919a, 'Descriptions', in Russell, B. 1919, *Introduction to Mathematical Philosophy*, George Allen and Unwin Ltd., London, p. 167-80, in Martinich, A.P. (ed.) 2001, *The Philosophy of Language*, 4<sup>th</sup> ed., Oxford University Press, Oxford, p. 221-7.
- 1919b, 'On Propositions: What They Are and How They Mean', in Marsh, R.B. (ed.) 1956, *Logic and Knowledge: Essays 1901-1950*, George Allen & Unwin Ltd, London, p. 285-320.
- Saul, J. 2005, 'Feminist Philosophy of Language', Stanford Encyclopedia of Philosophy, <<http://plato.stanford.edu/entries/feminism-language/>>
- Szabó, Z.G. 2005, 'Introduction', in Szabó, Z.G. (ed.) 2005, *Semantics versus Pragmatics*, Oxford University Press, Oxford, p. 1-14.
- Searle, J. 1969, *Speech Acts*, Cambridge University Press, New York.
- 1979, *Expression and Meaning*, Cambridge University Press, New York.
- Searle, J. and D. Vanderveken 1985, *Foundations of Illocutionary Logic*, Cambridge University Press, Cambridge.
- Sellars, W. 1991, 'Empiricism and the Philosophy of Mind', in Sellars, W. 1991, *Science, Perception and Reality*, Ridgeview Publishing Company, Atascadero, p. 127-96.
- Shope, R.K. 1983, *The Analysis of Knowing: A Decade of Research*, Princeton University Press, Princeton.
- Soames, S. 1999, *Understanding Truth*, Oxford University Press, Oxford.
- 2002, *Beyond Rigidity: The Unfinished Semantic Agenda of Naming and Necessity*, Oxford University Press, Oxford.
- Solstad, D. 1992, *Elleve roman, bok atten*, Forlaget Oktober, Oslo.
- Sophocles, 1992, 'Kong Oedipus', in Vandvik, E. 1992, *Greske tragediar*, Det Norske Samlaget, Oslo, p. 176-221.
- Sosa, E. forthcoming, 'Timothy Williamson's *Knowledge and Its Limits*', in Greenough, P. and D. Pritchard (eds.) forthcoming, *Williamson and His Critics*, Oxford University Press, Oxford.
- Stanley, J. 2005, *Knowledge and Practical Interest*, Oxford University Press, Oxford.
- Steup, M. 2006, 'The Analysis of Knowledge', Stanford Encyclopedia of Philosophy, <<http://plato.stanford.edu/entries/knowledge-analysis/>>
- Stine, G. 1976, 'Skepticism, Relevant Alternatives, and Deductive Closure', *Philosophical Studies* 29, D. Reidel Publishing Company, Dordrecht-Holland, p. 249-61, in DeRose, K. and T.A. Warfield (eds.) 1999, *Scepticism: A Contemporary Reader*, Oxford University Press, Oxford, p. 145-55.
- Tarski, A. 1944, 'The Semantic Conception of Truth and the Foundations of Semantics', *Philosophy and Phenomenological Research* 4, p. 341-75, in Martinich, A.P. (ed.) 2001, *The Philosophy of Language*, 4<sup>th</sup> ed., Oxford University Press, Oxford, p. 69-91.
- Unger, P. 1984, 'Philosophical Relativity', in Unger, P. 1984, *Philosophical Relativity*, University of Minnesota Press, Minneapolis, in DeRose, K. and T.A. Warfield (eds.) 1999, *Scepticism: A Contemporary Reader*, Oxford University Press, Oxford, p. 243-71.
- 2002 [1975], *Ignorance: A Case for Scepticism*, Oxford University Press, Oxford.
- Vogel, J. 2005, 'Can Skepticism be Refuted? The Refutation of Skepticism', in Steup, M. and E. Sosa (eds.), *Contemporary Debates in Epistemology*, Blackwell Publishing, Oxford, p. 72-84.
- Williamson, T. 1996, *Vagueness*, Routledge: Taylor & Francis Group, London.
- 2000a, *Knowledge and Its Limits*, Oxford University Press, Oxford.
- 2000b, 'Scepticism and Evidence', *Philosophy and Phenomenological Research*, 60, p. 613-28.
- 2004, 'Past the Linguistic Turn?', in Leiter, B. (ed.) 2004, *The Future for Philosophy*, Oxford University Press, Oxford, p. 106-28.
- 2005a, 'Contextualism, Subject-Sensitive Invariantism, and Knowledge of Knowledge', in *The Philosophical Quarterly*, vol. 55, nr 219, Blackwell Publishing, Oxford, p. 213-35.
- 2005b, 'Knowledge and Scepticism', in Jackson, F. and M. Smith (eds.) 2005b, *The Oxford Handbook of Contemporary Philosophy*, Oxford University Press, Oxford, p. 681-700.
- 2005c, 'Knowledge, Context, and the Agent's Point of View', in Peter, G. and G. Preyer (eds.) 2005, *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, Oxford University Press, Oxford, p. 91-114.



-----, 2006, 'Can Cognition be Factorised into Internal and External Components?', in Abrahamsen, A., W. Bechtel, R. Mathews and R. Stainton (eds.) 2006, *Contemporary Debates in Cognitive Science*, Blackwell Publishing, Oxford, p. 291-306.