

Anti-Individualism, Self-Knowledge, and Epistemic Possibility:
Further Reflections on a Puzzle About Doubt¹

Gary Ebbs
Draft 1/28/11

Forthcoming in A. Hatzimoysis, ed., *Self-Knowledge* (Oxford University Press, 2011)

In Ebbs 2003 I argue that there is no way to reconcile anti-individualism and minimal self-knowledge—the familiar fact (yet to be clarified) that we each know without empirical investigation what thoughts our own utterances express—with what I take to be the standard analysis of epistemic possibility. The standard analysis of epistemic possibility implies that if we suspend all of our empirical beliefs, we can each make sense of actually being in any possible world that is “subjectively indistinguishable” to us (in a sense I will explain) from the actual world. A central premise of my argument is that *a person can entertain or express a thought only if she has some idea of what that thought is*. But for any person there is a wide range of utterances such that if anti-individualism is true, then if she suspends all of her empirical beliefs, she does not have any idea what thoughts, if any, those utterances express. Among these are utterances she makes when she takes herself to be describing possible worlds that are “subjectively indistinguishable” to her from the actual world. If she does not have any idea what thoughts, if any, those utterances express, then, given the italicized premise, she cannot entertain or express thoughts about possible worlds that are “subjectively indistinguishable” to her from the actual world, contrary to the standard analysis of epistemic possibility.

My goal in this paper is to strengthen my case against the standard analysis of epistemic possibility by presenting a related but different argument for rejecting it. Unlike the argument just sketched, the new one does not directly depend on the claim that for any

person there is a wide range of utterances such that if anti-individualism is true, then if she suspends all of her empirical beliefs, she does not have any idea what thoughts, if any, those utterances express. The new argument stresses, instead, that we cannot have any justification for believing that a given description of a possible world w supports anti-individualism unless we accept some substantive empirical beliefs our acceptance of which implies that we cannot make sense of being in w .

Relative to other philosophical problems, the problem with the standard analysis of epistemic possibility that I aim to expose here is especially difficult to discern. The problem can remain hidden from us in much the way that the impossibility of the scene depicted in M. C. Escher's drawing "Waterfall" remains hidden from a viewer who focuses only on isolated parts of it, not on how those parts are related to each other in the whole drawing. To appreciate Escher's "Waterfall" one must see both how its parts can each seem to depict possible scenes when viewed in isolation from the whole, and why they fail to fit together into a whole that depicts a possible scene. Similarly, to understand and solve the problem that I aim to expose, one must see both how anti-individualism, self-knowledge, and the standard analysis of epistemic possibility can each seem plausible when considered in its own, and why anti-individualism and self-knowledge are together incompatible with the standard analysis of epistemic possibility. My new argument against the standard analysis of epistemic possibility is part of my explanation of why anti-individualism and self-knowledge are together incompatible with the standard analysis of epistemic possibility.

As preparation for presenting my new argument (in sections I-VII) I first clarify the preliminary observations about anti-individualism and self-knowledge that generate the

problem and then present the problem itself, making explicit how each step relies on the preliminary observations. I also state and justify a key conditional that I left implicit in Ebbs 2003, and address several objections to the reasoning that generates the problem. With this extensive background in place, (in sections VIII-IX) I present the new argument as a reply to a tempting but ultimately confused objection to my recommendation that we reject the standard analysis of epistemic possibility.

I. Background Assumptions and the Metalinguistic Mirroring Principle

My preliminary characterization of minimal self-knowledge (as the familiar fact that we each know without empirical investigation what thoughts our own utterances express) immediately raises two questions:

- (1) What is it to know what thoughts our own utterances express?
- (2) What is it to know this “without empirical investigation”?

Let us consider question (2) first. Many philosophers take “without empirical investigation” to be synonymous with “a priori,” and assume that *a priori* knowledge (or belief) is knowledge (or belief) based *solely* on reasoning or introspecting (or both).

I argue in Ebbs 2003, however, that if anti-individualism is true, there is a sense in which we each know without empirical investigation what thoughts our own utterances express, even though we cannot know what thoughts our own utterances express *a priori*, where “a priori” is understood to mean “solely by reasoning or introspecting”.

My argument begins with reflections about the Twin Earth thought experiments (adapted from Putnam 1975 and Burge 1979) that are widely taken to establish anti-

individualism. These thought experiments are rooted in our practice of taking fellow English speakers' words at face value. Suppose there is a person named Oscar who is at least minimally competent in the use of the English word 'water' but who neither accepts nor rejects the sentence 'Water is H₂O'. If Oscar sincerely utters the sentence 'water is a liquid at room temperature', then without hesitation other English speakers, ourselves included, will take his words at face value—i.e. they will take him to have expressed the thought that *water is a liquid at room temperature*. Now suppose that on Twin Earth, which is just like Earth except that wherever there is water on earth, there is twin water, a liquid with a chemical structure very different from that of water, there lives a physical, phenomenological, and behavioral *twin* of Oscar, call him Twin Oscar, who is competent in Twin English, the language spoken on Twin Earth. If Twin Oscar sincerely utters the sentence 'water is a liquid at room temperature', then without hesitation other Twin English speakers will take his words at face value—i.e. they will take him to have expressed the thought (translated into English) that *twin water is a liquid at room temperature*.

If we trust our practice of taking other speakers' words at face value, then these observations support anti-individualism. Philosophers typically make these observations without conducting any special empirical inquiries. It may therefore seem that the observations are based on reasoning and introspecting alone, hence independent of our beliefs about how things are in the actual world. When we make the observations, however, we presuppose a range of beliefs about the actual world, including our belief that water is H₂O, that we live on Earth, and that Twin Earth is different from Earth in having twin water, a substance different from water, in its rivers, lakes, and oceans. If we were to

reject these presuppositions, thereby leaving open the possibility that Earth is Twin Earth and water is twin water, then we would have no reason to conclude that the thought that Oscar expresses when he utters his sentence ‘water is a liquid at room temperature’ is different from the thought that Twin Oscar expresses when he (Twin Oscar) utters his sentence ‘water is a liquid at room temperature’. Hence if we were to reject the presuppositions, we could not take the thought experiment to support anti-individualism.

One might think that we could support anti-individualism just by reasoning or introspecting if we simply *stipulate* that Earth is not Twin Earth and water is not twin water. (See, for instance, McLaughlin 2004: 433.) But, for reasons I explain in Ebbs 1996 and Ebbs 2001, anti-individualism itself implies that these are not claims whose truth values can be settled by stipulation. We may accept them provisionally, to see what follows from them, for instance, but in doing so, we would be accepting assumptions that could turn out to be false.

To support anti-individualism, then, we must presuppose beliefs that cannot be known just by reasoning or introspecting. We nevertheless take for granted that *without going through any special empirical investigation of our own*, we are entitled to accept the background beliefs on which our arguments for anti-individualism rely. If for some reason we conclude that we are not entitled to accept those background beliefs without going through any special empirical investigation of our own, then we cannot support anti-individualism, and hence we should not accept it. These reflections support the first of the two main Background Assumptions for the puzzle:

Background Assumption 1: We can support anti-individualism *without going through any special empirical investigation of our own*, but we can do so only if we presuppose beliefs that cannot be known just by reasoning and introspecting.

This crucial Background Assumption is related to minimal self-knowledge. We typically trust our practice of taking each other's words at face value unless there are good reasons in a given context for not doing so. For reasons I have just briefly sketched, if we take this practice to provide us with our best grip on what speakers of a natural language say and believe, we commit ourselves to anti-individualism. A crucial related observation is that our practice of taking each other's words at face value goes hand-in-hand with attributions of minimal competence in the use of words. The requirements for minimal competence in a given context are typically not high: as Putnam famously observed, for instance, a person who cannot tell the difference between an elm tree and a beech tree may still be regarded as minimally competent in the use of the words 'elm' and 'beech', so that her utterances of sentences containing those words are taken at face value and evaluated accordingly. In each context, there are nevertheless some minimal requirements for competence in the use of words. A person who applies 'beech' only to beaches, for instance, will not be regarded as competent in the use of the word 'beech'. When such a person utters the sentence "Elms are not beeches," we will not take him to have said that elms are not beeches, but instead, perhaps, that elms are not beaches. In short, to take a speaker's words at face value is to take her to have minimal competence in the uses of those words. And we will not take her to have minimal competence in the uses of those words if we do not also take her to express some of her substantive beliefs by using those words. (Ebbs 2003: 146-147.)

Consider how these reflections bear on question (1) above—the question “What is it to know what thoughts our own utterances express?” The natural answer, given anti-individualism, is that to have minimal self-knowledge of what thoughts one’s current utterances express is just to use one’s words to express those thoughts.² When we take another’s words at face value we thereby credit her with having minimal self-knowledge. One might think that minimal self-knowledge must be based solely on reasoning or introspecting —i.e. that “without empirical investigation” in our characterization of minimal self-knowledge means “solely by reasoning or introspecting”. But if anti-individualism is true, minimal self-knowledge cannot be independent of empirical beliefs in this sense, since it requires minimal competence, which we acquire by accepting testimony from others. As J. L. Austin observed, “Reliance on the authority of others is fundamental . . . for corroboration and for the correctness of our own use of words, which we learn from others” (Austin 1979: 83, fn1). When we learn the word “elm”, for instance, we are told that elms are *trees*. Either we trust what we are told, or we seek another authority, perhaps a book or an expert on trees. Our competence in ‘elm’ is not rooted in our own empirical investigation into whether the word ‘elm’ applies to trees, but in our acceptance of what we are told or what we read about elms. We are typically entitled to accept what we are told or what we read about elms unless we have some special reason not to do so (Ebbs 2003: 148).

Hence if we adopt anti-individualism for the reasons sketched above, we should not take “without empirical investigation” in our characterization of minimal self-knowledge to mean “solely by reasoning or introspecting”. Instead, we should take “without empirical investigation” in our characterization of minimal self-knowledge to mean “without going

through any special empirical investigation of one's own." When minimal self-knowledge is understood in this way, the observations in the previous two paragraphs support

Background Assumption 2: If anti-individualism is true, then to have minimal self-knowledge is to know what thoughts one's utterances express without going through any special empirical investigation of one's own, while presupposing beliefs one acquired by testimony from others.

By displaying links between anti-individualism, minimal competence, and minimal self-knowledge, this Background Assumption supports and clarifies the truism that *to express a thought, one must have some idea of what that thought is.*

There is one more point about minimal self-knowledge that is crucial to the puzzle I'll present below. If, as I just argued, to have minimal self-knowledge of what thoughts one's current utterances express is to be able to use one's words to express those thoughts, then a person may have minimal self-knowledge even if she does not actually entertain any metalinguistic beliefs about what thoughts her utterances express. We are nevertheless each in a position to entertain metalinguistic beliefs about what thoughts our utterances sentences express. Let's say that a disquotational metalinguistic belief is a belief expressed by a sincere utterance of a sentence of the form

(D) My utterances of ' _____ ' in the language I now actually speak as I now actually use it express the thought that _____,³

where the blanks are filled by non-indexical declarative sentences that the utterer has practical knowledge of how to use. (Indexical sentences introduce complications that we need not go into here.) Anyone who knows how to use the sentence 'Water is a liquid at room temperature' and understands disquotation is in a position to assert "My utterances of

‘Water is a liquid at room temperature’ in the language I now actually speak as I now actually use it express the thought that water is a liquid at room temperature.” When I assert this sentence, for instance, I thereby express my disquotational meta-linguistic belief that my utterances of ‘Water is a liquid at room temperature’ in the language I now actually speak as I now actually use it express the thought that water is a liquid at room temperature.

I propose that we think of the connection between minimal self-knowledge and disquotational metalinguistic beliefs as follows. To have minimal self-knowledge of what thoughts one’s current utterances express is just to be able to use one’s words to express those thoughts. Hence if we are able to use the words in the sentence ‘Water is a liquid at room temperature’ to express the thought that water is a liquid at room temperature and we understand disquotation, we are justified in taking our present utterances of those words to express the thought that water is a liquid at room temperature (Ebbs 1996: 509-510). This reasoning supports the following conditional:

If S has practical knowledge of how to use the words in a given non-indexical declarative sentence *s* and S understands disquotation, then S is justified in accepting the corresponding disquotational metalinguistic belief about what thoughts S’s utterances of *s* express.

The converse is also true:

If S is justified in accepting the corresponding disquotational metalinguistic belief about what thoughts S’s utterances of a given non-indexical declarative sentence *s* express, then S has practical knowledge of how to use the words in *s* and S understands disquotation.

The conjunction of these claims is what I shall call the

Metalinguistic Mirroring Principle: S has practical knowledge of how to use the words in a given non-indexical declarative sentence *s* and S understands disquotation if and only if S is justified in accepting the corresponding disquotational metalinguistic belief about what thoughts S's utterances of *s* express.

II. Subjectively equivalent worlds and substantive beliefs

Twin Earth thought experiments like the one described above presuppose that some person on Earth and his twin on Twin Earth are in the same subjective state, in some philosophically important senses of “same” and “subjective”, even if the thoughts they express by their respective utterances are different. More generally, Twin Earth thought experiments suggests that for each person we can describe *subjectively equivalent worlds* in which everything that is relevant to the person's subjective assessment of her situation seems the same to her as it does in the actual world, but her social or physical environments may be different from her social and physical environment in the actual world. Moreover, the worlds described in Twin Earth thought experiments are by assumption metaphysically possible. If we model our notion of the subjectively equivalent world on the worlds described in Twin Earth thought experiments, then, we may conclude that subjectively equivalent worlds are metaphysically possible worlds, and that one of the subjectively equivalent worlds in question is the actual world.⁴

In what sense, however, are the worlds subjectively equivalent? In previous writings, I did not try to articulate the notion of subjective equivalence in other terms, such

as sameness of phenomenology, on the assumption that philosophers may agree that two worlds *are* subjectively equivalent, yet disagree about what *makes* them subjectively equivalent. My hesitation on this point bothers Brian McLaughlin, who seeks a more informative definition of “subjectively equivalent world”. He proposes that we define it as follows:

A [metaphysically] possible world *w* is a subjectively equivalent world for a subject *S* if and only if in *w* *S* has exactly the same history of intrinsic physical states and exactly the same history of phenomenal experiences that *S* has in the actual world.

(McLaughlin 2004, p. 428)

One problem with this definition is that phenomenal experiences may themselves be individuated anti-individualistically, so that, for instance, Oscar has the visual experience as of seeing that there is a glass of *water* in front of him while simultaneously his twin on Twin Earth has the visual experience as of seeing (as we would express it) that there is a glass of *twin water* in front of him. Yet even if we suppose that Oscar’s and Twin Oscar’s visual experiences are individuated anti-individualistically, we may still believe that Oscar’s visual experience is in some epistemically important sense *subjectively indistinguishable* from, hence subjectively equivalent to, Twin Oscar’s visual experience.

This is not merely a logically possible position: according to Michael Martin’s disjunctive account of perceptual experience, for instance, “the phenomenal character of two experiences can be different even while one of them is indistinguishable from the other.”

(Martin 2006, 366) For this reason, I prefer the following definition:

w is a subjectively equivalent world for a subject S if and only if (a) w is a metaphysically possible world, and (b) in w S has the same history of intrinsic physical states that S has in the actual world.⁵

I assume that philosophers who differ about how to understand phenomenal experiences can agree on this definition of “ w is a subjectively equivalent world for a subject S ”, and take sameness of intrinsic physical states to imply subjective equivalence.

We can now define “the proposition that p is substantive for S ” in terms of subjectively equivalent worlds, as follows:

The proposition that p is *substantive* for S if and only if p is true in some of S 's subjectively equivalent worlds and p is false in some of S 's subjectively equivalent worlds.

With this definition in place, we can specify a weak *sufficient* condition for affirming that a person S holds a substantive belief that p :

Substantive Belief (sufficient condition): If subject S accepts the proposition that p , p is substantive for S , and S 's acceptance of p is sincere, then S holds a *substantive* belief that p .

It is notoriously difficult to state a plausible necessary condition for holding a belief.

However, the following very weak necessary condition should be uncontroversial:

Substantive Belief (necessary condition): If a subject S holds any substantive beliefs, then it is not the case that S neither (sincerely) accepts nor (sincerely) rejects any of the substantive propositions that S considers.

The contrapositive of this necessary condition is “if S neither (sincerely) accepts nor (sincerely) rejects any of the substantive propositions that S considers, then S holds no substantive beliefs.”

Although we don't have exhaustive necessary and sufficient conditions for being a substantive belief, the above definition of “the statement that p is substantive for S” and Substantive Belief (sufficient condition) together strongly suggest that a substantive belief is one that cannot be known just by reasoning and introspecting. With this in mind, we can rewrite Background Assumptions 1 and 2 as follows:

Background Assumption 1': We can support anti-individualism without going through any special empirical investigation of our own, but we can do so only if we presuppose some substantive beliefs.

Background Assumption 2': If anti-individualism is true, then one can know what thoughts one's utterances express without going through any special empirical investigation of our own, but one can know this in that way only if one presupposes some substantive beliefs.

From Background Assumptions 1' and 2' it follows that if we accept anti-individualism, but suspend or ignore all of our substantive beliefs, we will no longer be able to use our words in discourse in the sense described above, and will hence will lose our minimal self-knowledge. In particular, if I suspend or ignore all of my substantive beliefs, I will be unable simultaneously to describe some proper subset of my subjectively equivalent worlds, such as the set $\{x: x$ a subjectively equivalent world in which I was born, raised, and now live on Twin Earth, where my word 'water' refers to twin water, not water}, and wonder whether the actual world is a member of that set, because I will not be minimally

competent to use any of the words some of which I need to be minimally competent to use if I am to describe those sets of subjectively equivalent worlds.

While I stand behind this consequence of Background Assumption 1' and 2', I realize that many philosophers who accept anti-individualism are inclined to reject it.⁶ To find an argument that may persuade such philosophers, I propose that we consider the kinds of thoughts we need to be in a position to entertain and justify if we are to take ourselves to describe some proper subset of our subjectively equivalent worlds without just stipulating that one of the worlds in the set is actual. I shall argue below (in section VIII) that a careful consideration of these kinds of thoughts supports what I call the Substantive Belief Condition, which sheds light on the puzzle without directly prejudging the question of whether we have self-knowledge if we suspend all of our substantive beliefs. Between now and then (in sections III-VII) I shall complete my clarification of the preliminary reasoning on which a proper understanding of the puzzle relies, sketch the puzzle itself, and discuss several different responses to it.

III. Epistemic possibility and subjectively equivalent worlds

Given Background Assumptions 1' and 2', how should we understand the relationship between what is epistemically possible for a given subject S, on the one hand, and worlds that are subjectively equivalent for S, on the other? Suppose, provisionally, that we accept:

Epistemic Possibility (EP): It is epistemically possible for a given subject S (at time t) that S is actually in world w if and only if (a) w is a subjectively equivalent world for S , and (b) S can make sense of w 's being actual (at t).

Here we explain epistemic possibility in terms of a subject's being able to make sense of actually being in some subjectively equivalent world. As I noted above, a subjectively equivalent world is also a metaphysically possible world. It follows from EP, then, that it is epistemically possible for a given subject S (at a given time t) that S is actually in world w only if w is a metaphysically possible world.

For this reason, EP is at best only an analysis of *empirical* epistemic possibility. It does not work for the sort of epistemic possibility relevant to mathematics and logic. As Kripke 1980 observes, for instance, it is epistemically possible for us that Goldbach's conjecture (that every even number greater than two is the sum of two primes) is true, and epistemically possible for us that Goldbach's conjecture is false. If Goldbach's conjecture is false, then there is no metaphysically possible world in which Goldbach's conjecture is true, so we cannot analyze the epistemic possibility that Goldbach's conjecture is true in terms of the existence of a metaphysically possible world in which it is true—by hypothesis, there is no such world. Similarly, if Goldbach's conjecture is true, we cannot analyze the epistemic possibility that Goldbach's conjecture is false in terms of the existence of a metaphysically possible world in which it is false. Hence we must take EP to be at best an analysis of empirical epistemic possibility. To know when to apply EP to a given epistemic possibility, we therefore need to know whether that possibility is empirical. In order to examine the role of EP in the Standard Analysis of Epistemic Possibility, I assume,

provisionally, that we can know whether not a given epistemic possibility is empirical just by reasoning or introspecting.⁷

Not all accounts of epistemic possibility imply that if it is epistemically possible for a given subject *S* (at a given time *t*) that *S* is actually in world *w*, then *w* is metaphysically possible. On some accounts of epistemic possibility, it may be epistemically possible for a given subject *S* (at *t*) that *S* is actually in world *w*, where *w* is *not* a metaphysically possible world. (For an example, see Soames 2005: 81-83.) Unlike EP, such accounts of epistemic possibility do not explain epistemic possibility in terms of subjectively equivalent (hence metaphysically possible) worlds and are therefore not limited by their explanatory strategy to being analyses of empirical epistemic possibilities. I shall endorse this less limited explanatory strategy below.

When considering the consequences of anti-individualism, however, it is very tempting to analyze epistemic possibility in terms of subjectively equivalent worlds. Those who are persuaded by the standard arguments for anti-individualism typically also take for granted Kripke's view of metaphysical necessity. And Kripke himself proposed an analysis like EP in order to deal with a puzzling consequence of his view of metaphysical necessity. In Ebbs 2003, I explain this puzzling consequence as follows:

In Kripke's view, if Hesperus is identical to Phosphorus, then there is no possible world in which Hesperus is not identical to Phosphorus. Prior to our discovery that Hesperus is identical to Phosphorus, we assumed that 'Hesperus is identical to Phosphorus' may actually be false. Hesperus is in fact identical to Phosphorus, however, and so, by Kripke's theory, Hesperus is *necessarily* identical with

Phosphorus: we can't express our prior assumption by saying it could have turned out that Hesperus is not identical to Phosphorus. (Ebbs 2003:149)

To express our prior assumption that 'Hesperus is identical to Phosphorus' may actually be false, Kripke suggested that "given the evidence that someone has antecedent to his empirical investigation, he can be placed in a sense in exactly the same situation, that is a qualitatively identical epistemic situation, and call two heavenly bodies 'Hesperus' and 'Phosphorus', without their being identical." (Kripke 1980: 104) Kripke observes that prior to our empirical investigation, we could not discriminate between these worlds on the basis of our evidence, and suggests we may actually have been in a world in which 'Hesperus is identical to Phosphorus' is not true. Thus Kripke apparently endorses EP for empirical epistemic possibilities. He assumes that prior to our discovery that Hesperus is identical to Phosphorus, it was epistemically possible for us that we were actually in a subjectively equivalent world in which our sincere utterances of 'Hesperus is identical to Phosphorus' are not true.⁸

Can we apply EP to the subjectively equivalent worlds that figure in Twin Earth thought experiments? It may seem that the answer is "yes". Recall, for instance, the Twin Earth thought experiment involving Oscar and Twin Oscar described above. Many philosophers who subscribe to EP maintain that Oscar can *make sense* of actually being in a subjectively equivalent world in which his word 'water' refers to twin water, not water. What is it for Oscar to make sense of this? It is very tempting to suppose that Oscar can make sense of it simply by *picturing himself* existing in a subjectively equivalent world in which his word 'water' refers to twin water, not water. By hypothesis a subjectively equivalent world is subjectively indiscriminable from the actual world, in some

philosophically interesting sense of “subjectively indiscriminable”, so there’s no barrier from the relevant “subjective” perspective to imagining oneself in any one such world. In addition, to give content to the thought that one is picturing oneself in one of a proper subset of one’s subjectively equivalent worlds, such as the set of subjectively equivalent worlds in which one’s word ‘water’ refers to twin water, not water, it seems enough to add a caption to the picture that tells us what it pictures. It seems enough to add, for instance, “I am now picturing myself in a subjectively world in which my word ‘water’ refers to twin water, not water,” or “I am now picturing myself in a subjectively equivalent world in which my names ‘Hesperus’ and ‘Phosphorus’ denote different planets.” In this way, I suggest, it is tempting to conjoin EP with

Picturing: If S can picture herself existing in a world w that S has reason to believe is subjectively equivalent for S, then S can make sense of w ’s being actual.

By conjoining EP with Picturing we arrive at what I call the Standard Analysis of Epistemic Possibility.

IV. The Justification Constraint

To get a preliminary sense of why the Standard Analysis of Epistemic Possibility is problematic, consider the difference between the claim that *it is epistemically possible for S (at time t) that she is actually in world w* , on the one hand, and the claim that *S (at t) cannot tell by reasoning or introspecting alone whether or not she is actually in world w* , on the other. An account of epistemic possibility is satisfactory only if it enables us to distinguish

between these claims. The best way to do so, I believe, is to adopt the following constraint on a satisfactory account of epistemic possibility:

Justification Constraint: It is epistemically possible for S (at a given time t) that she is actually in world w only if S's actually being in w is compatible with what S (at t) takes herself to know.⁹

For instance, if I take myself to know that my word 'water' refers to H₂O, then I cannot without contradiction picture myself in a subjectively equivalent world in which my word 'water' does not refer to H₂O. For to picture myself in such a world is to suppose that the world is actual, hence that I am in a world in which my word 'water' refers to H₂O and my word 'water' does not refer to H₂O. But there is no such subjectively equivalent world. All subjectively equivalent worlds are metaphysically possible, hence also logically possible. What I take myself to know at any given time t is limited—there is always *more than one* subjectively equivalent world that I can describe which is not *ruled out* by what I take myself to know at t , and hence there is always more than one subjectively equivalent world that I can make sense of being in at t . Suppose at t I take myself to know that my word 'water' refers to H₂O, but I do not take myself to know that Hesperus is the second planet from the Sun. Then I *cannot* picture myself in a subjectively equivalent world in which my word 'water' does not refer to H₂O, but I *can* picture myself in a subjectively equivalent world in which Hesperus is not the second planet from the Sun.

These observations show, paradoxically, that we can (at a given time t) *picture* ourselves existing in *any one* of our subjectively equivalent worlds only if we suppose (at t) that we do not know—in particular, that we are not epistemically justified in accepting or entitled to accept—*any* of our substantive beliefs. We might be tempted to see this as

compatible with the rationalist idea that if we restrict ourselves to what we can know just by reasoning or introspecting, we do not have any substantive knowledge. But the supposition that if we restrict ourselves to what we can know just by reasoning or introspecting, we can picture ourselves existing in any one of our subjectively equivalent worlds is in tension with a fundamental consequence of Background Assumptions 1' and 2' stressed above: if I suspend or ignore all of my substantive beliefs, hence ignore or suspend all of what I take to be my substantive knowledge, I will be unable simultaneously to describe some proper subset of my subjectively equivalent worlds, such as the set $\{x: x$ a subjectively equivalent world in which I was born, raised, and now live on Twin Earth, where my word 'water' refers to twin water, not water $\}$, and wonder whether the actual world is a member of that set, because I will not be minimally competent to use any of the words some of which I need to be minimally competent to use if I am to describe those possibilities.

Proponents of the Standard Analysis of Epistemic Possibility (including Brian McLaughlin, as we shall see) typically do not see the difficulty just described. They suppose that we can always suspend or disregard our substantive beliefs, yet still use our utterances to express thoughts, and know what thoughts our utterances express. Hence they assume that if we restrict ourselves to what we can know just by reasoning or introspecting, then it is epistemically possible for us that we are actually in any one of our subjectively equivalent worlds. The puzzle that I shall now at last present and clarify is designed to show that this assumption conflicts with anti-individualism and minimal self-knowledge.

V. The puzzle

The puzzle consists of four statements that we are each apparently justified in accepting without special empirical investigation. I shall present the reasoning in the first person, with the understanding that we can each accept each step in the reasoning, as applied to ourselves. In the first step of the puzzle, I accept the Standard Analysis of Epistemic Possibility (i.e. the conjunction of EP and Picturing), and assume (contrary to what I argued above) that I know what thoughts my utterances express even if I suspend or ignore all of my substantive beliefs. Moreover, I feel confident that I can picture myself existing in any of my subjectively equivalent worlds. From EP and Picturing, I infer

(1) I may actually be in any one of my subjectively equivalent worlds.

If I accept (1), then by the justification constraint I must suspend my substantive beliefs and infer that I have no substantive knowledge. The remaining steps of the puzzle are designed to show that I cannot accept this consequence of (1), and hence cannot accept (1) itself, if I also accept anti-individualism and take myself to have minimal self-knowledge.

In the second step of the puzzle, I recall that by Background Assumption 2', without going through any special empirical investigation of my own, yet while presupposing beliefs I acquired by testimony from others, I know what thoughts my own utterances express, in the sense that I have practical knowledge of how to use the words that occur in the sentences I utter. In particular, I know what thoughts my own utterances of the sentence 'Water is a liquid at room temperature' express, in the sense that I have practical knowledge of how to use the words that occur in that sentence. I also understand disquotation. I can therefore formulate the metalinguistic sentence

- (2) My utterances of ‘Water is a liquid at room temperature’ in the language I now actually speak as I now actually use it express the thought that water is a liquid at room temperature,

By the Metalinguistic Mirroring Principle, I am justified in accepting (2) without going through any special empirical investigation.

In the third step of the puzzle, I realize that

- (3) For every subjectively equivalent world w , my utterances of ‘Water is a liquid at room temperature’ in the language I now speak in w as I now use it in w express the thought that water is a liquid at room temperature.

is related to (1) and (2) as follows:

Key Conditional (KC): If I am justified in accepting (1) and (2) without special empirical investigation, then I am justified in accepting (3) without special empirical investigation.¹⁰

The argument for KC is as follows.

Argument for KC. Suppose (toward contradiction) that I am justified without special empirical investigation in accepting (1) and (2), but not (3). Assume also that bivalence holds for (3). Then if I am not justified in *accepting* (3) without special empirical investigation, I am not justified in *denying* the *negation* of (3) without special empirical investigation. The negation of (3) is

- (4) For some subjectively equivalent world w , my utterances of ‘Water is a liquid at room temperature’ in the language I now speak in w as I now use it in w do not express the thought that water is a liquid at room temperature.

Hence I am not justified in denying (4) without special empirical investigation, or, in other words, (4) is epistemically compatible with what I am justified in believing without special empirical investigation. Let w be an existential instantiation of (4). Then in the language I now speak in w as I now use it in w my utterances of ‘Water is a liquid at room temperature’ do not express the thought that water is a liquid at room temperature. By supposition I am justified in accepting (1) without special empirical investigation. But (1) implies that w may be actual, and hence that I may *actually* speak the language that I speak in w . Therefore, if (4) is compatible with what I am justified in believing without special empirical investigation, then so is the supposition that I am actually in w and hence actually speak the language that I speak in w . But if I *am* in w , then in the language I actually

speak as I actually use it, my utterances of 'Water is a liquid at room temperature' do not express the thought that water is a liquid at room temperature. By assumption I am not justified in believing without special empirical investigation that I am not in *w*. Hence

- (a) I am not justified in believing without special empirical investigation that my utterances of 'Water is a liquid at room temperature' in the language I now actually speak as I now actually use it express the thought that water is a liquid at room temperature.

But, by supposition I am justified in accepting (2) without special empirical investigation. And according to (2), my utterances of 'Water is a liquid at room temperature' in the language I now speak as I now use it express the thought that water is a liquid at room temperature. Hence

- (b) I am justified in believing without special empirical investigation that my utterances of 'Water is a liquid at room temperature' in the language I now actually speak as I now actually use it express the thought that water is a liquid at room temperature.

The conjunction of (a) and (b) is a contradiction. Hence, by *reductio ad absurdum*, if I am justified in accepting (1) and (2) without special empirical investigation, I am justified in rejecting (4), i.e. the negation of (3), without special empirical investigation. But if I am justified in rejecting the negation of (3) without special empirical investigation, then I am justified in accepting (3) without special empirical investigation. We may therefore conclude that if I am justified in accepting (1) and (2) without special empirical investigation, then I am justified in accepting (3) without special empirical investigation.

We now arrive at the fourth and final step of the puzzle: By Background assumption 1' and Twin Earth thought experiments of the sort described above, I am justified in accepting (4) without special empirical investigation. But, as we have seen, (4) is just the negation of (3). The conjunction of (3) and (4) is a contradiction.

To avoid this contradiction I must give up either (3) or (4). Hence, given KC, to avoid this contradiction I must give up either (1), (2), or (4). As we have seen, my reasons for accepting (1) without special empirical investigation seem plausible when they are considered in isolation from my reasons for accepting (2) and (4) without special empirical investigation. Unlike my reasons for accepting (1) without special empirical investigation, however, my reasons for accepting (2) and (4) without special

empirical investigation go hand-in-hand with my acceptance of anti-individualism. I and other anti-individualists should therefore accept (2) and (4), but reject (1). Since the Standard Analysis of Epistemic Possibility implies that I am justified in accepting (1) without special empirical investigation, I and other anti-individualists should therefore reject the Standard Analysis of Epistemic Possibility.

VI. An Objection to the Argument for the Key Conditional

Anti-individualists who wish to defend the Standard Analysis of Epistemic Possibility must reject some part of the reasoning just sketched. Let us begin with the Argument for KC. It is sound? Consider the following central part of it:

Let w be an existential instantiation of (4). Then in the language I now speak in w as I now actually use it in w my utterances of ‘Water is a liquid at room temperature’ do not express the thought that water is a liquid at room temperature. By supposition I am justified in accepting (1) without special empirical investigation. But (1) implies that w may be actual, and hence that I may actually speak the language that I speak in w . Therefore, if (4) is compatible with what I am justified in believing without special empirical investigation, then so is the supposition that I am actually in w and hence actually speak the language that I speak in w .

One might be tempted to raise the following objection to this part of the argument for KC.

Objection to the Argument for KC. With or without special empirical investigation, it is not an epistemic possibility for me that I am in some world other than the world I am actually in. Hence (1) should be understood to hold for a given subjectively equivalent world w only in a context in which we describe w without stipulating that w is not actual. But if I suppose that w is an existential instantiation of (4), then in effect I describe w as some world other than the world that I am actually in. If I suppose that w is an existential instantiation of (4), then I can tell from my description of w that it is not an epistemic possibility for me that I am in w and hence actually speak the language that I speak in w . Hence if I suppose that w is an existential instantiation of (4), then, under that description, (1) does not apply to

w , and I cannot infer from (1) that I may actually be in w , contrary to the Argument for KC.¹¹

I accept the first two sentences of this objection. Hence in my view the plausibility of the objection rests on the plausibility of the claim that if I suppose that w is an existential instantiation of (4), then I describe w as some world other than the world that I am actually in. Let us now examine this claim.

To begin with, consider why (1) should be understood to hold for a given subjectively equivalent world w only in a context in which we describe w without stipulating that w is not actual. Consider the following reasoning:

Suppose w is a non-actual subjectively equivalent world. Since it is not an epistemic possibility for me that I am right now in some world other than the world I am actually in, I know without special empirical investigation that I am in the actual world, and that w is a non-actual. I therefore know without special empirical investigation that I am not in w .

There is nothing wrong with this reasoning, but it does not undermine (1). The idea behind (1) is that given an *informative* description of a subjectively equivalent world w that does *not* explicitly build in the assumption that w non-actual, one cannot tell without special empirical investigation whether one is actually in w . More precisely, if (1) is true and w is one of my subjectively equivalent worlds, then there is what I shall call a *canonical* description D of w —i.e. an *informative* description of w that does not explicitly beg the question of whether w is the actual world.

The objection to the Argument for KC raises the question: “Do I provide a canonical description of w when I describe w as an instantiation of (4), hence as a subjectively equivalent world w such that my utterances of ‘Water is a liquid at room

temperature' in the language I now speak in w as I now actually use it in w do not express the thought that water is a liquid at room temperature?" The answer is "yes."

To see why, it helps to compare this description of w with a description of a subjectively equivalent w' world that relies on indexicals. For example, suppose I am kidnapped, blindfolded and moved some unknown distance from where I was when I could last locate myself. In these circumstances, of course, I don't know where I am. Still, I can reason as follows:

I am here now. Let w' be one of my subjectively equivalent worlds in which I am one mile north of here. Then I know without special empirical investigation that I am not actually in w' .

This reasoning is correct, but only because the description D' of w' that it relies on, namely, " w' is one of my subjectively equivalent worlds in which I am one mile north of here", is not canonical. For I get no information about where I am by sincerely uttering "I am here now," and then reasoning that I am not anywhere else, including one mile north of where I am, in part because this description builds in the supposition that w' is not the actual world.

One crucial difference between D' and my description D of a subjectively equivalent world w such that my utterances of 'Water is a liquid at room temperature' in the language I now speak in w as I now actually use it in w do not express the thought that water is a liquid at room temperature is that, unlike D' , which relies on the indexical expression "here" to specify w' , D does *not* rely on indexical expressions to specify w . The English sentence "Water is a liquid at room temperature" contains no indexicals, and its occurrence in D does not express an indexical thought.¹² D is also unlike the description D'' of a subjectively equivalent world w'' such that my utterances of 'Water is a liquid at

room temperature' in the language I now speak in w'' as I now actually use it in w'' express something other than what they actually now express. For D'' does *not* inform me that my utterances of 'Water is a liquid at room temperature' in the language I now speak in w'' as I now actually use it in w'' do not express the thought that water is a liquid at room temperature.

One might think that the actual world is the only subjectively equivalent world w such that my utterances of "Water is a liquid at room temperature" in the language I now speak in w as I now actually use it in w express the thought that water is a liquid at room temperature. It might then be easy to conflate the supposition that w is a subjectively equivalent world such that my utterances of "Water is a liquid at room temperature" in the language I now speak in w as I now actually use it in w do not express the thought that water is a liquid at room temperature with the empty, uninformative supposition that w is a non-actual subjectively equivalent world, because one takes the two sets

$W = \{w: w \text{ is a subjectively equivalent world and my utterances of "Water is a liquid at room temperature" in the language I now speak in } w \text{ as I now actually use it in } w \text{ do not express the thought that water is a liquid at room temperature}\}$

and

$X = \{w: w \text{ is a non-actual subjectively equivalent world}\}$

to be identical. There are two problems with this. First, even if W and X *were* identical, one could not know this just by reflecting on their descriptions. Second, W and X are not, in fact, identical: there are non-actual subjectively equivalent worlds for me such that my utterances of "Water is a liquid at room temperature" in the language I now speak in these worlds as I now actually use it in these worlds express the thought

that water is a liquid at room temperature. For instance, consider the subjectively equivalent world in which the truth value of “there is an even number of books on my bookshelf right now” is different from what it actually is. I have no idea whether there is an even number of books on my bookshelf. If there is, then a world in which there isn’t, because one book was removed from the shelf without my noticing, would be subjectively indistinguishable to me. But in some of these worlds my utterances of “Water is a liquid at room temperature” in the language I now speak as I now actually use it express the thought that water is a liquid at room temperature. Hence there are some non-actual subjectively equivalent worlds for me such that my utterances of “Water is a liquid at room temperature” in the language I now speak in these worlds as I now actually use it in these worlds express the thought that water is a liquid at room temperature.

Finally, while W is identical with Y , where

$Y = \{w: w \text{ is a subjectively equivalent world and my utterances of “Water is a liquid at room temperature” in the language I now speak in } w \text{ as I now actually use it in } w \text{ express something other than what they now (i. e. actually) express}\}$

we cannot know that W is identical with Y without relying on (2), which, as I argued above, we cannot know without presupposing substantive beliefs. One might be tempted to deny this, if one equated

My utterances of “Water is a liquid at room temperature” in the language I now speak in w as I now actually use it in w express what they actually now express with

My utterances of “Water is a liquid at room temperature” in the language I now speak in w as I now actually use it in w express the thought that water is a liquid at room temperature.

But these two sentences do not say the same thing—the first is empty, the second, informative.

I conclude that the description D of w in the argument for KC is *canonical* in the sense defined above: it is an informative description of w that does not explicitly beg the question of whether w is the actual world. This undermines the Objection to the Argument for KC.

VII. Can an anti-individualist reject (4)?

One might accept the Argument for KC, yet reject my conclusion that anti-individualists should reject (1), either because one believes that anti-individualists are not entitled to accept (2) without special empirical inquiry, or that anti-individualists are not entitled to accept (4) without special empirical inquiry. To believe that anti-individualists are not entitled to accept (2) without special empirical inquiry, however, one would have to give up either the Mirroring Principle or the supposition that even if anti-individualism is true, we are each entitled to suppose that we have minimal self-knowledge in the sense explained above. Neither of these is open to serious doubt. A marginally more attractive option is to believe that anti-individualists are not entitled to accept (4) without special empirical inquiry. In support of this option, one might reason as follows.

A standard Twin Earth thought experiment shows that there is at least one world w such that w is subjectively equivalent for me and my utterances of ‘Water is a liquid

at room temperature' in the language I now speak in *w* as I now actually use it in *w* express the thought that *twin water* (in particular, XYZ) is a liquid at room temperature. From this I cannot infer (4) unless I also have grounds for supposing that water is not identical to twin water. But I cannot know that water is not identical to twin water if I suspend all my empirical beliefs. Hence if I rely on reasoning and introspecting alone, I cannot support (4).

One can refuse to affirm (4) for this reason, yet still accept (1) and (2). By the Argument for KC, one is thereby committed to accepting (3). But if, for the reasons just explained, an anti-individualist is not committed to (4), then an anti-individualist who also accepts (1) and (2) is not committed to the contradictory conjunction of (3) and (4), and so one does not have to reject (1) to accept anti-individualism and (2).

The trouble with this way of trying to solve the puzzle is that it comes at the cost of undermining one's grounds for accepting anti-individualism. As I explained in section I, if I cannot rely on the Twin Thought experiments to support (4), then for all I can tell from the thought experiments that I accept, my utterances of any given sentence express the *same* thought in all my subjectively equivalent worlds. Hence I have no grounds for concluding that what a person believes and thinks is not settled by his linguistic dispositions, internal physical states, or phenomenal experiences, described independently of his social and physical environment. In short, I have no grounds for accepting anti-individualism. But the puzzle arises precisely *because* I take myself to have justification for accepting anti-individualism, and hence, in particular, for accepting (4). Moreover, for reasons I explained above, my justification for accepting anti-individualism cannot be based only on reasoning or introspection. We must presuppose some substantive beliefs if we are to take ourselves to be describing alternative subjectively equivalent worlds and asking what our words refer to in such worlds. These substantive beliefs provide us with

grounds for accepting (4). Yet this is not to say that we need to engage in special empirical inquiry to accept (4). These points should by now be familiar. They are summed up in

Background Assumption 1': We can support anti-individualism without going through any special empirical investigation of our own, but we can do so only if we presuppose some substantive beliefs.

In short, contrary to the suggestion we have been investigating, we *are* entitled to accept (4) without special empirical inquiry.

VIII. McLaughlin's attempt to solve the puzzle without rejecting (1)

Is there any other way to try to resist my conclusion (in §V) that anti-individualists must accept (2) and (4), but reject (1)? Brian McLaughlin argues that one can accept (1), (2), and (4) without special empirical investigation, yet reject (3). If the Argument for the KC is sound, McLaughlin's reasoning cannot be correct. Nevertheless, it is instructive to see where it goes wrong. McLaughlin's strategy is to distinguish between metaphysical possibility and epistemic possibility, as characterized by the Standard Analysis of Epistemic Possibility. He reasons as follows:

If anti-individualism is true, then there are subjectively equivalent worlds for me in which I use 'water' to refer to a chemical substance XYZ. And if [(1)] is true, then I may be in a subjectively equivalent world in which I use 'water' to refer to a chemical substance XYZ. But ... I cannot know independently of empirical investigation whether water is XYZ. Since XYZ is by stipulation not H₂O, if water = H₂O, then it is *metaphysically impossible* that water is XYZ. But according to the

standard conception of epistemic possibility, it is nonetheless *epistemically possible* that water is XYZ. (McLaughlin 2004: 433)

There are three problems with the reasoning in this passage. First, in the penultimate sentence, McLaughlin asserts that XYZ is by stipulation not H₂O. He apparently takes this to imply that I can know solely by reasoning or introspecting that XYZ is not H₂O. As I noted in §I, however, if anti-individualism is true, then if I suspend all of my substantive beliefs, I have no grounds for believing that XYZ is not H₂O. Moreover, I cannot regard this belief as true by stipulation, because I can describe circumstances in which it's false.¹³ Second, since I cannot know a priori that XYZ is not H₂O, I cannot know a priori that "if water = H₂O, then it is metaphysically impossible that water is XYZ."

Third, and most important, if we suppose (following McLaughlin, who agrees with Kripke) that

If water is H₂O, and H₂O is not XYZ, then it is metaphysically necessary that water is H₂O.

then, since every subjectively equivalent world is metaphysically possible, there is no subjectively equivalent world in which water is XYZ. Hence given EP, if water is H₂O, and H₂O is not XYZ, it is not epistemically possible for me that water is XYZ, contrary to what McLaughlin says.

McLaughlin will likely reply that this last objection misses his point, which is that even if there is no subjectively equivalent world in which water is XYZ, there is a subjectively equivalent world in which I use 'water' to refer to XYZ. (This reading is supported by the second sentence of the passage just quoted.) And, according to the Standard Analysis of Epistemic Possibility, it is epistemically possible for me that I am in

such a subjectively equivalent world if I can *picture* myself actually existing in it. Hence if we understand “it is epistemically possible for me that water is XYZ” to mean “it is epistemically possible for me that I am in a subjectively equivalent world in which I use ‘water’ to refer to XYZ” then it seems we can allow that it is epistemically possible for me that water is XYZ.

This reply overlooks the fact that subjectively equivalent worlds are metaphysically possible worlds. And Kripke and Putnam have shown that to judge which worlds are metaphysically possible worlds, we must presuppose substantive beliefs. Is there a metaphysically possible world in which water is not H₂O? The answer depends on whether water is H₂O in the actual world. For if it is, according to Kripke and Putnam, then there is no metaphysically possible world in which water is not H₂O. “According to this way of looking at things,” Scott Soames explains, “in order to find out whether certain things are true with respect to **all** possible states of the world, and other things are true with respect to **no** possible states of the world, we sometimes must **first** find out what is true with respect to the actual state of the world.” (Soames 2005: 82) This conception of metaphysical possibility is built into the usual understanding of the possible worlds used in the Twin Earth thought experiments that support anti-individualism. And again, these worlds are supposed to be subjectively equivalent worlds for us. We know a priori that the actual world is metaphysically possible. But if *w* is one of a proper subset of subjectively equivalent worlds that we describe without stipulating that it is the actual world, then, for the reasons just explained, we cannot know a priori that *w* is metaphysically possible, and hence we cannot know that *w* is one of our subjectively equivalent worlds. For instance, consider the set

$Z = \{w: w \text{ a subjectively equivalent world in which I was born, raised, and now live on Twin Earth, where my word 'water' refers to twin water, not water}\}.$

To know that Z is non-empty I would need to know, of some world w , that w is a member of Z . (The ordinary procedures for constructing Twin Earth thought experiments convince us that Z is non-empty by convincing us, of some world w canonically described in the experiment, that w is a member of Z .) But unless I simply stipulate that w is the actual world, I cannot know solely by reasoning or introspecting that w is metaphysically possible. Since a subjectively equivalent world is a metaphysically possible world, it follows that unless I simply stipulate that w is the actual world, I cannot know solely by reasoning or introspecting that w is one of my subjectively equivalent worlds. But to be a member of Z , w must be one of my subjectively equivalent worlds. Hence I cannot know solely by reasoning or introspecting that w is a member of Z . This reasoning supports the

Substantive Belief Requirement: If w is one of a proper subset of subjectively equivalent worlds that S describes without simply stipulating that w is the actual world, then S has reason to believe that w is subjectively equivalent for him (and hence metaphysically possible) only if S holds some substantive beliefs.

From the Substantive Belief Requirement it follows immediately that if S restricts himself to what he can know solely by reasoning or introspecting, and w is one of a proper subset of worlds that S describes without simply stipulating that w is the actual world, then S has no reason to believe that w is one of his subjectively equivalent worlds.

More generally, by the Substantive Belief Requirement, if w is one of a proper subset of subjectively equivalent worlds that S describes without simply stipulating that w is the actual world, then S has reason to believe that w is subjectively equivalent for him

(and hence metaphysically possible) only if S holds some substantive beliefs. But either (a) I have reason to believe that w is a member of Z, or (b) I do not. If (a), then, by the Substantive Belief Requirement, I hold some substantive beliefs. Hence I have reason to believe that w is a member of Z only if I do *not* restrict myself to what I can know solely by reasoning or introspecting. But then, by the Justification Constraint I can infer that there are subjectively equivalent worlds, including all the worlds in Z, that I cannot coherently suppose that I am in. Hence, contrary to what McLaughlin assumes, I cannot accept (1). If (b), then, by EP, I have no grounds for believing that it is epistemically possible for me that I am in Z. Either way, then, given the Standard Analysis of Epistemic Possibility, I have no grounds for believing that it is epistemically possible for me that I am in Z. I must therefore reject (1).

IX. Substantive beliefs and self-undermining doubts

To support my recommendation that we reject (1), in Ebbs 2003 I present an argument designed to show that if I accept anti-individualism, I cannot make sense of the supposition that I am actually in a subjectively equivalent world w in which I was born, raised, and now live on Twin Earth, where my word ‘water’ refers to twin water. The argument runs as follows¹⁴:

- (i) I am now using this sentence to express the thought that water is a liquid at room temperature.
- (ii) If I were actually in a subjectively equivalent world w in which I was born, raised, and now live on Twin Earth, where my word ‘water’ refers to twin water, not water,

I could not use sentence (i) to express the thought that water is a liquid at room temperature.

- (iii) Therefore, I am not actually in a subjectively equivalent world w in which I was born, raised, and now live on Twin Earth, where my word 'water' refers to twin water, not water.

If we accept anti-individualism, then by Background Assumptions 1' and 2' we can each take ourselves to know the premises of this argument without special empirical investigation. Since the argument is valid, we can each derive (iii) without special empirical investigation. But this conflicts with the supposition that without special empirical investigation, we can accept (1), the claim that we may actually be in *any one* of our subjectively equivalent worlds. Hence anyone who understands the argument is committed to rejecting (1) without special empirical investigation.

McLaughlin grants that this argument is sound, but claims that it does not undermine the Standard Analysis of Epistemic Possibility. He assumes that according to that analysis, we are each justified in accepting (1) just by reasoning or introspecting. To state his objection to the argument, he distinguishes between a *weak* and a *strong* sense of "knowable independently of empirical investigation." The weak sense is the one I recommend: "knowable independently of any special empirical investigation of one's own, but not independently of all one's substantive beliefs." The strong sense is "knowable just by reasoning or introspecting." McLaughlin's central objection to the argument (i)-(iii) is as follows:

In [the] weak sense of "knowable independently of empirical investigation", the standard analysis of epistemic possibility does not entail that if p is epistemically

possible, then p is not knowable independently of empirical investigation. A proponent of the standard analysis of epistemic possibility thus need not deny that one can know independently of empirical investigation in Ebbs's weak sense that one is not actually in the kind of world in which one was born, raised, and now lives on Twin Earth, even if the proponent holds that it is epistemically possible that one is actually in the kind of world in which one was born, raised, and now lives on Twin Earth. So, even if Ebbs were right that via the argument from [(i)] and [(ii)] to [(iii)] one can know in this weak sense independently of empirical investigation that one is not in the kind of world in which one was born, raised, and now lives on Twin Earth, it would not follow that the standard analysis of epistemic possibility is mistaken.¹⁵

This objection is correct so far as it goes: argument (i)-(iii) relies on the weak sense of “knowable independently of empirical investigation”, and the Standard Analysis of Epistemic Possibility is widely *assumed* to hold for the strong sense of “knowable independently of empirical investigation”. But the objection overlooks Background Assumptions 1' and 2'. As I emphasized above, from Background Assumptions 1' and 2' it follows that if I suspend or ignore all of my substantive beliefs, I will be unable simultaneously to describe some proper subset of my subjectively equivalent worlds, such as Z , and wonder whether the actual world is a member of that set, because I will not be minimally competent to use any of the words some subset of which I need to be minimally competent to use if I am to describe those sets of subjectively equivalent worlds. This is just another application of the truism that to express a thought, one must have some idea of what that thought is.

In the previous section we saw that there is another route to a similar conclusion. According to the Substantive Belief Requirement, if w is one of a proper subset of subjectively equivalent worlds that S describes without simply stipulating that w is the actual world, then S has reason to believe that w is subjectively equivalent for him (and hence metaphysically possible) only if S holds some substantive beliefs. Hence if I restrict myself to what I know solely by reasoning and introspecting, then I have no justification for believing that it is epistemically possible for me that the actual world is a member of the set Z . Contrary to what McLaughlin supposes in his defense of the Standard Analysis of Epistemic Possibility, given that Analysis, I am *not* justified in claiming to know in the strong sense of “without empirical investigation” that it is epistemically possible that I am actually in a subjectively equivalent world that is in the set Z .

In short, we are not justified in accepting (1) “without empirical investigation,” on the strong reading of that phrase. And both the argument of section V and the supplementary argument (i)-(iii) of this section show that we are not justified in accepting (1) “without empirical investigation,” on the weak reading of that phrase, either. We may therefore conclude that we are not justified in accepting (1) “without empirical investigation,” on *either* the strong *or* the weak reading of that phrase. The argument of section V and the supplementary argument (i)-(iii) each establish this conclusion if we presuppose Background Assumptions 1' and 2', which together yield an anti-individualistic interpretation of the truism that to express a thought, one must have some idea of what that thought is. They also establish this conclusion if we accept the Substantive Belief Requirement. As we saw in section VIII, the Substantive Belief Requirement is a natural consequence of the conception of metaphysical possibility that goes hand in hand with

Twin Earth thought experiments that justify our acceptance of (2) and (4) without special empirical investigation. I conclude that if the Twin Earth thought experiments justify our acceptance of (2) and (4) without special empirical investigation for the reasons given above, we must reject (1), and, with it, the Standard Analysis of Epistemic Possibility.¹⁶

Works Cited

- Austin, J. L. 1979: "Other Minds," in J. O. Urmson and G. J. Warnock, eds., *Philosophical Papers*, third edition (Oxford: Oxford University Press), pp. 76-116.
- Burge, T. 1979: "Individualism and the Mental", in Peter A. French, Theodore E. Uehling, and Howard K. Wettstein, *Midwest Studies in Philosophy* 4, Minneapolis: University of Minnesota Press, pp. 73-122.
- 1982: "Other Bodies," in *Thought and Object*, A. Woodfield, ed. (London: Oxford University Press).
- Ebbs, G. 1996, "Can We Take Our Words at Face Value?" *Philosophy and Phenomenological Research*, Volume 56, Number 3 (1996): 499-530.
- 2001, "Is Skepticism about Self-Knowledge Coherent?" *Philosophical Studies* 105 (2001): 43-58.
- 2003, "A Puzzle About Doubt," in Susana Nuccetelli, ed., *New Essays on Semantic Externalism and Self-Knowledge* (MIT Press).
- 2005, "Why Skepticism About Self-Knowledge is Self-Undermining," in *Analysis*, Volume 65, Number 3 (2005).
- Kripke, S. 1980: *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Martin, M.G.F. 2006, "On Being Alienated," in Tamar Szabò Gendler and John Hawthorne, eds., *Perceptual Experience* (Oxford: Clarendon Press 2006).
- McLaughlin, B. 2004, "Anti-Individualism and Minimal Self-Knowledge: A Dissolution of Ebbs's Puzzle" in Richard Shantz, ed., *The Externalist Challenge* (Berlin: Walter de Gruyter).
- Putnam, H. 1975: "The Meaning of 'Meaning'," in Hilary Putnam, *Mind, Language, and Reality: Philosophical Papers*, Vol. 2. Cambridge: Cambridge University Press, 1975, pp. 215-271.
- 1988: *Representation and Reality* Cambridge, MA: MIT Press.
- Soames, S. 2005, *Reference and Description: The Case Against Two-Dimensionalism* (Princeton: Princeton University Press).

¹Thanks to Brian McLaughlin, whose criticisms in McLaughlin 2004 of my arguments in Ebbs 2003 prompted me to write this paper, and to Kate Abramson and Adam Leite for their helpful comments on a previous draft.

²For a parallel point about what it is to know the *meanings* of one's own words, see Putnam 1988: 32.

³In Ebbs 2003, I expressed (D) less cautiously with the sentence

(*) My utterances of ‘_____’ express the thought that _____.
Instantiating this sentence to ‘Water is liquid at room temperature’ yields

(**) My utterances of ‘Water is liquid at room temperature’ express the thought that water is liquid at room temperature.

In almost all philosophical discussions of disquotation, a formulation such as (**) is understood to be restricted to my utterances of ‘Water is liquid at room temperature’ in the language I now speak as I now use it. Unfortunately, however, (**) does not *rule out* the interpretation according to which (**) “concerns all my utterances, *whether past, present, or future*, of ‘Water is liquid at room temperature.’” (McLaughlin 2004: 432.) And, of course, I cannot know without special empirical investigation (or perhaps even *with* it) that all my past or future utterances of ‘Water is liquid at room temperature’ express the thought that Water is liquid at room temperature. By formulating (*) as (D), we rule out the interpretation just described.

⁴McLaughlin argues that “if the actual world were not one of one’s subjectively possible worlds, then it would be trivially true that one can know just by reasoning or introspecting--without empirical investigation--which of one’s subjectively equivalent worlds one is actually in. The reason would be that one is not actually in any of them.” (McLaughlin 2004, p. 428) But by McLaughlin’s own legalistic standards, this reasoning is mistaken. For it may be that, on some understanding of the phrase “subjectively possible,” the actual world is not *in fact* one of one’s “subjectively possible” worlds, and yet one could not know this just by reasoning or introspecting. It is a trivial *logical* truth, of course, that if one can know just by reasoning or introspecting that the actual world is not one of one’s subjectively possible worlds, then one can know just by reasoning or introspecting that the actual world is not one of one’s subjectively possible worlds. But this is presumably not what McLaughlin has in mind.

⁵McLaughlin is puzzled by my preference for this definition. See McLaughlin 2004, note 7. He apparently overlooks the possibility of combining anti-individualism with disjunctivism about perceptual experience.

⁶The main reason, I think, is that in practice it is virtually impossible to suspend or ignore one’s substantive beliefs. Philosophers who accept anti-individualism but still feel that they understand what thoughts their sentences express when they (try to) suspend or ignore all their substantive beliefs in fact always *presuppose* some of those beliefs. Their *feeling* that they understand what thoughts their sentences express therefore doesn’t support their *assumption* that if they were somehow to succeed in suspending or ignoring all their substantive beliefs, they would still be in a position to understand what thoughts their sentences express.

⁷Given our definition of “subjectively equivalent world,” EP implies that certain “possibilities” that some philosophers *take* to be empirical epistemic possibilities for us, such as the “possibility” that we are disembodied spirits with no physical bodies at all, are *not* empirical epistemic possibilities for us. In order to examine the role of EP in the Standard Analysis of Epistemic Possibility, I provisionally accept this consequence of EP as well.

⁸We cannot describe this as a subjectively equivalent world in which Hesperus is not identical to Phosphorus, since by assumption is not metaphysically possible that Hesperus is not identical to Phosphorus. This snag in Kripke’s analysis leads Scott Soames to reject EP and to conclude that for Kripke “what is epistemically possible is not always metaphysically possible.” (Soames 2005: 82) I, too, think we should reject EP. But I think EP is an initially appealing analysis that many philosophers of language at least tacitly endorse.

⁹This constraint is compatible with the sorts of epistemic possibilities characteristic of arguments for radical skepticism, if we grant that to accept those radical possibilities as epistemic possibilities for us is to agree with the skeptic that our ordinary assumptions about what we know are false.

¹⁰In Ebbs 2003 I left it as an exercise for the reader to supply KC and see the argument for it, because I thought they were obvious. That was a mistake. Brian McLaughlin (and, no doubt, other thoughtful readers) saw that I needed something like the KC, but did not see how to formulate it or why it is true.

¹¹Adam Leite raised a version of this objection in response to an earlier draft of this paper.

¹²For a detailed explanation and defense of this observation, see Burge 1982.

¹³For a description of how we might discover that water is twin water, despite our assumption that it is not, see Ebbs 1996: fn32. A similar thought experiment could be devised to make sense of the remote possibility that XYZ is H₂O, despite our assumption that it is not. Neither of these skeptical scenarios requires that we suspend all of our substantive beliefs.

¹⁴Ebbs 2003: 156. I made a few changes that clarify the premises. First, I replaced the definite article “the” by “a”, since there is more than one subjectively equivalent world of the kind in question. Second, I added the clause “where my word ‘water’ refers to twin water, not water,” so that the kind of world in question is stated in the premises themselves.

¹⁵McLaughlin 2004: 435.

¹⁶To reject the Standard Analysis of Epistemic Possibility is to reject either EP or Picturing (or both). In Ebbs 2003 I recommended that we reject Picturing, and even suggested (in Ebbs 2003: fn 32) that we reject EP as well. Soames 2005 argues in detail that we should reject EP. As I mentioned above, according to Soames, it may be epistemically possible for a given subject S (at a given time t) that S is actually in world *w*, where *w* is *not* a metaphysically possible world. (Soames 2005: 82). But if *w* is not a metaphysically possible world then *w* is not a subjectively equivalent world for S. Hence it may be epistemically possible for a given subject S (at t) that S is actually in world *w*, even if condition (a) of EP is not satisfied. To accept this, however, is *not* to say that we can have minimal self- knowledge if we suspend all of our substantive beliefs. By Background Assumption 2', we are still committed to the truism that *to express a thought, one must have some idea of what that thought is*. Moreover, the argument (i)-(iii) shows that some *apparently* coherent skeptical suppositions, such as the supposition that I am actually in a subjectively equivalent world in which I was born, raised, and now live on Twin Earth, are self-undermining. In Ebbs 2003 I propose that “we take ourselves to *know* that we are not in some specified subjectively equivalent world if from our perspective, any attempt to specify how we may actually be in that subjectively equivalent world is self-undermining.” (Ebbs 2003: 160; for an explanation and defense of the proposal, see Ebbs 2003: 158-162.)