

Subjective externalism

Article (Accepted Version)

Sawyer, Sarah (2018) Subjective externalism. *Theoria*, 84 (1). pp. 4-22. ISSN 1755-2567

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/71315/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

Subjective Externalism

by

SARAH SAWYER

University of Sussex

1. Introduction

Individualism is the view that all of an individual's representational mental kinds are constitutively independent of any relation to a wider reality. *Anti-individualism* maintains that many representational mental states and events are constitutively what they are partly by virtue of relations between the individual in those states and a wider reality. (Burge, 2007: 3)

Individualism and anti-individualism are in fundamental disagreement about the constitutive nature of representational mental kinds. As a result, they have mutually inconsistent implications concerning the attribution of thoughts to individuals.

Individualism implies that intrinsic duplicates necessarily instantiate the very same representational mental kinds—that intrinsic duplicates necessarily have the same thoughts. This is because an individual's relations to her environment play no constitutive role in individuating her representational mental kinds, and hence the thoughts an individual has are constitutively what they are independently of any relations she has to her particular environment. Anti-individualism, in contrast, implies that intrinsic

duplicates will not necessarily instantiate the same representational mental kinds, since which representational mental kinds an individual instantiates depends partly, but essentially, on relations between her and her environment. As such, anti-individualism implies that intrinsic duplicates may well instantiate different representational mental kinds—may well have different thoughts—in virtue of being related to different objective properties in their respective environments.

Individualism and anti-individualism are mutually inconsistent theories. An individual's representational mental states are either constitutively independent of any relations to a wider reality, or they are not. In this paper I aim to draw out the motivations that lie behind each theory, and to formulate, for each theory, a constraint on the correct attribution of thought content to which it is committed. What we will find is that the individualist is motivated by the idea that an adequate theory of thought content must respect a specific internal, subjective constraint on the attribution of thought content, while the anti-individualist is motivated by the idea that an adequate theory of thought content must respect a specific external, objective constraint on the attribution of thought content.¹ I will argue that while the theories are themselves inconsistent, the specific constraints to which each is committed are not. This realisation opens up the conceptual space for a different kind of position—a position according to which the correct attribution of thought content respects *both* the internal, subjective constraint that motivates individualism *and* the external, objective constraint that motivates anti-individualism. For reasons that will become clear as we proceed, I call the resulting view

¹ It is not part of my thesis that the motivations or constraints I identify are explicitly stated by proponents of the views I discuss. Rather, I maintain that the views are in fact committed to those constraints and are best understood as being motivated by the relevant commitments.

‘subjective externalism’, thereby distinguishing it from the traditional anti-individualist view, which, for the purposes of the paper and in order to highlight the contrast, I label ‘essentialist externalism’. The resulting view is a form of anti-individualism, but one that respects the internal, subjective constraint on the attribution of thought content that motivates individualism.

The structure of the paper is as follows. In section 2 I present the traditional anti-individualist position and state the external constraint on the attribution of thought content that motivates it. In section 3 I present the individualist position and state the internal constraint on the attribution of thought content that motivates it. In section 4 I introduce subjective externalism. I conclude briefly in section 5.

2. An External Constraint

According to the anti-individualist, non-representational, causal relations to objective properties in one’s environment partly determine what one can represent in thought. Non-representational relations to different objective properties allow for different representational mental kinds. Anti-individualism is notoriously supported by so-called Twin Earth thought experiments.² These invite reflection on counterfactual scenarios in which an individual’s intrinsic make-up is hypothesized to remain constant while the wider reality in which she is embedded is hypothesized to differ. According to the anti-individualist, the hypothesized differences in the individual’s wider environment would

² See Putnam (1973) for the original ‘Twin Earth’ thought experiment as applied to linguistic meaning. See also Kripke (1972). See Burge (1979), (1982) and (1986) for three different applications to mental kinds. Putnam accepts the application to mental kinds in his (1996).

result in corresponding differences in the individual's thoughts. Take the following example. Suppose a subject S is related in the right kind of non-representational way to silver—she has a silver bracelet, her mother has a silver tray, and so on. As such, she acquires the concept *silver*, and comes to believe that silver must be polished to prevent it from tarnishing. Now consider a counterfactual scenario in which there is no silver; a counterfactual scenario in which S's bracelet and her mother's tray are not made of silver, but are instead made of a hypothetical look-alike metal which S would be unable to distinguish either practically or theoretically from silver. Let's call the look-alike metal in the counterfactual scenario 'twilver'. In the counterfactual scenario, S is related in the right kind of non-representational way not to silver, but to twilver—she has a twilver bracelet, her mother has a twilver tray, and so on. This makes it plausible to think that whereas S acquires the concept *silver* and comes to believe that silver must be polished to prevent it from tarnishing, counterfactual S acquires the concept *twilver*, and comes to believe that twilver must be polished to prevent it from tarnishing. The difference in representational content between the belief S has and the belief S would have lies in relations to her wider environment—non-representational relations to silver and twilver respectively.

The Twin Earth thought experiments make it clear that anti-individualism is committed to an external, objective constraint on the attribution of thought content, which can be stated as follows:

(EC) The correct attribution of thought content to an individual is essentially constrained by her non-representational relations to objective properties in her wider reality.

Although the example given relates to a so-called ‘natural kind’, anti-individualism in all its varieties is committed to (EC).³ But the Twin Earth thought experiments suggest that anti-individualism is committed in addition to the claim that the content of an individual’s mental state can outstrip her individual perceptual discriminatory capacities. In the example given above, S’s representational mental kinds are supposed to differ from counterfactual S’s representational mental kinds even though silver and twilver are such that S’s individual perceptual discriminatory capacities are insufficient to distinguish between them.⁴ The point generalises. The content of an individual’s mental state is typically taken, by the anti-individualist, to depend constitutively not only on relations between her and certain objective properties in her environment, but on relations between her and certain objective properties in her environment to which she may be, as I will put it, ‘perspectivally blind’.

³ For a more detailed discussion of different kinds of anti-individualism, see Sawyer (2011).

⁴ I take discriminatory capacities generally to be contrastive, and am hence sympathetic with certain reasons for thinking a subject’s individual discriminatory capacities are context-relative. The contrastive, or context-relative nature of discriminatory capacities is particularly important for debates in epistemology. However, in the debate over the individuation of representational content, both individualists and anti-individualists agree that it is absolute discriminability rather than relative discriminability that is at issue. As such, for present purposes I set discussions of the contrastive, or context-relative nature of discriminability to one side. For a contrastive account of the propositional attitudes based on a contrastive account of empirical knowledge, and hence in part on the contrastive nature of discriminability, see Sawyer (2014).

Perspectival blindness is a notion that goes beyond the mere fallibility of our perceptual systems. Both the individualist and the anti-individualist acknowledge that our individual perceptual discriminatory capacities are prone to error. Mistaking a shadow for a spider, a child's laugh for a cry, or a glass of water for a glass of gin, are errors of a familiar kind. But they are errors that concern particular instances of properties on particular occasions, and, crucially, concern properties which the subject is in general able to identify by means of her current, individual perceptual discriminatory capacities. On closer inspection, it can quickly become clear that there was no spider, no cry, and no gin. Perspectival blindness, in contrast, concerns properties which the subject is not able to identify by means of her current, individual perceptual discriminatory capacities.⁵ I characterise perspectival blindness as follows:

(PB) A property P is one to which a subject S is perspectivally blind at time t if and only if there is a possible property P' distinct from P, such that S does not have the individual perceptual discriminatory capacity at t to distinguish any instance of P from any instance of P'.

Perspectival blindness is a relation between a subject and a property at a time, and so the phenomenon of perspectival blindness is subject- and time-relative. A subject who is

⁵ The notion of a subject's individual discriminatory capacities may be thought problematic for other reasons. For example, Hellie (2005) argues that discriminability raises phenomenal Sorites problems, and Hawthorne (2007) argues that perceptual discrimination is unstable because it relies on cognition, which is subject to interference. However, the relevant assumption in the present context is that a pair of twins, S and counterfactual S, will have the same individual perceptual discriminatory capacities even if these

perspectively blind to a property at one point in time may not be perspectively blind to that property at a later time, since her individual perceptual discriminatory capacities may change in relevant ways. In addition, perspectival blindness is defined in terms of a subject's capacities at a time, and not in terms of the exercise of those capacities at that time. Thus a subject does not become perspectively blind to a property merely by closing her eyes, wearing ear-plugs, failing to pay close enough attention, or falling asleep, for instance. Both individualism and anti-individualism acknowledge the phenomenon of perspectival blindness. Both theories acknowledge that S, in the example above, is perspectively blind to the property of being silver. But anti-individualism maintains in addition that an individual's mental kinds, states and events depend constitutively on relations between her and certain objective properties in her environment to which she may be perspectively blind. That is, according to the anti-individualist, perspectival blindness presents no barrier to representation.

This commitment is evident in the three main varieties of anti-individualism. According to natural kind anti-individualism the content of certain of an individual's mental states depends constitutively on the essential properties of the natural kind those states represent, even if the subject cannot as an individual discriminate perceptually between instances of the relevant natural kind and qualitative doubles whose essential natures differ.⁶ According to social anti-individualism, the content of certain of an individual's mental states depends constitutively on the way in which members of her

capacities are subject to Sorites problems and certain kinds of instability. As such, although the issues merit further discussion, they fall beyond the scope of the current paper.

⁶ There has been a tendency in the literature to focus on microstructural kinds, but whether the essential nature of a natural kind is microstructural will depend on the kind in question.

linguistic community use words, even if the subject cannot as an individual discriminate perceptually between the practices of her own linguistic community and those of a distinct community whose linguistic practices do in fact differ.⁷ And, according to singular thought anti-individualism, the content of certain of an individual's mental states depends constitutively on the very identity of the objects they concern, even if the subject cannot as an individual discriminate perceptually between those objects and qualitative doubles whose identities differ.⁸ It is this dependence of thought on essential properties, societal norms, or numerical identity that marks out the traditional conception of anti-individualism. But essential properties, societal norms and numerical identity are not just objective properties in the subject's environment, they are objective properties to which the subject may be perspectively blind. Henceforth, I will refer to anti-individualism as traditionally conceived as 'essentialist externalism'. The label is perhaps less well-suited to social anti-individualism than to natural kind or singular anti-individualism, but the label is less significant than the generality it is intended to capture. It might also be called 'blind externalism', since the defining mark of essentialist externalism is its commitment to the claim that the content of a subject's thoughts depends constitutively on objective properties in her wider reality to which she may be perspectively blind.

3 An Internal Constraint

⁷ See Burge (1979), (1982) and (1986). I have deliberately avoided talk of deference to experts in the discussion of social anti-individualism because, despite its popularity in the literature, there is reason to think it cannot be a fundamental part of anti-individualism. See Sawyer (2018).

⁸ See Evans (1982) and McDowell (1977), (1984) and (1986).

According to the individualist, an individual's representational mental kinds are constitutively independent of her relations to a wider reality. There are different versions of individualism; but each one is committed to an internal, subjective constraint on the attribution of thought content, which can be stated as follows:

(IC) The correct attribution of thought content to an individual is essentially constrained by the local supervenience of her mental states and events on her individual perceptual discriminatory capacities.

Individualism is often defined in terms of its commitment to the claim that a subject's mental states and events supervene locally on her intrinsic physical states and events. There are two reasons why I have not put (IC) in these terms. First, individualism as a generic position is not committed to the claim that a subject has physical states and events.⁹ Second, given the aims of the paper, the primary question concerns the relation between a subject's mental states and events on the one hand and her individual perceptual discriminatory capacities on the other.

It follows from (IC) that subjects with the same individual perceptual discriminatory capacities should be attributed mental states and events with the same content. In the example given in the previous section, S and counterfactual S have the same individual perceptual discriminatory capacities; things would seem the same to S were she in either the actual scenario related to silver or the counterfactual scenario related to twilver; the differences between the actual and the counterfactual scenarios are differences to which S

is, as explained above, perspectively blind. According to the individualist, and in line with (IC), then, S and counterfactual S should be attributed thoughts with the same content.

The fact that the differences between the actual and the counterfactual scenarios are differences to which the subjects are perspectively blind is crucial here. In particular, since the property of being silver is a property to which S is perspectively blind, she should not, according to the individualist, be attributed the concept *silver*; and since the property of being twilver is a property to which counterfactual S is perspectively blind, she should not, according to the individualist, be attributed the concept *twilver*.

Generalising, a subject should not be attributed a concept that refers to a property to which she is perspectively blind. To attribute a subject such a concept would lead to violations of the local supervenience thesis stated in (IC), since it would then be possible for subjects with the same individual perceptual discriminatory capacities to have thoughts with different representational contents.

Individualism does, of course, allow that properties to which an individual is perspectively blind (neurophysiological properties, for instance) can make a constitutive difference to her mental kinds, states and events, but these properties will have to be intrinsic to her rather than objective properties in her wider environment, and they will not provide the representational content of her thoughts. As a result, they do not lead to violations of the local supervenience thesis stated in (IC), and hence do not provide a counterexample to the claim that individualism is committed to (IC). The individualist, then, is committed to the claim that objective properties to which an individual is

⁹ Admittedly this is currently the standard individualist view, but there are exceptions provided by, for

perspectivally blind cannot be represented by that individual in thought.¹⁰ Putting things this way makes it clear that (IC) constitutes a rejection of the defining mark of essentialist externalism given towards the end of the previous section. Whereas perspectival blindness presents no barrier to representation according to the traditional forms of anti-individualism, perspectival blindness does present a barrier to representation according to the individualist.

That (IC) represents a common commitment of the different versions of individualism can be seen by examining them briefly but systematically.¹¹ There are four main individualist positions. The first is a form of descriptivism, according to which a subject should be attributed (rather than, for example, the concept *silver* or the concept *twilver*) compositional concepts that meet two conditions. First, the compositional concepts must be composed of concepts which map directly on to the subject's individual perceptual discriminatory capacities, and hence refer to properties to which the subject is not perspectivally blind. This ensures that the compositional concepts will themselves supervene locally on the subject's individual perceptual discriminatory capacities and refer to properties to which the subject is not perspectivally blind. Second, the compositional concepts must explicitly encapsulate an individual's beliefs about the relevant subject matter. For example, the concept *S* possesses in the scenario described

instance, Chalmers (2003) and Descartes (1641).

¹⁰ Individualism is, of course, a thesis about what *determines* representational content. But for this very reason, it is also a thesis about *representational content*. This serves to connect standard interpretations of individualism with the claim that objective properties to which an individual is perspectivally blind cannot be represented by that individual in thought.

above might be thought of not as the natural kind concept *silver*, but as the descriptive concept *shiny metallic substance often used to make jewellery and trays and that needs to be polished to prevent it from tarnishing*. Treating concepts descriptively in this way allows subjects with the same individual perceptual discriminatory capacities to be attributed the very same concepts despite their relations to different environments. This is because the descriptions are designed specifically to capture the way things seem from the subject's perspective. Counterfactual S would, on this view, be attributed the very same descriptive concept as S, and the shared concept would correctly apply to silver, twilver, and any other substance that fits the description S and counterfactual S have in mind. This kind of individualism, then, is clearly committed to (IC).

The second individualist position also attributes concepts designed to capture the way things seem from the individual's perspective, but rejects the compositional account of concepts advocated by the descriptivist, and instead advocates an atomistic understanding of concepts.¹² On this atomistic view, the concept S possesses in the scenario described above is a concept that has in its extension silver, twilver and anything else that S cannot distinguish from them, just as it does on the descriptivist view. As such, from our theoretical perspective, S's concept can be identified descriptively as the concept that applies to shiny metallic substances often used to make jewellery and trays and that needs to be polished to prevent it from tarnishing. But, according to the atomistic view, the concept is not itself compositional or descriptive in form. To keep this fact firmly in

¹¹ My aim at this juncture is merely to characterize the various individualist positions and establish the common commitment to an internal, subjective constraint on the correct attribution of thought content. For a more detailed discussion of the positions, see Sawyer (2007) and (2011).

¹² For this view see Segal (2000).

mind, it is, perhaps, best to introduce a new term, such as ‘shmilver’, stipulatively to be a term that expresses the concept S possesses. The attribution of thought content to an individual preserves, on this view, the local supervenience of a subject’s mental states and events on her individual perceptual discriminatory capacities, and so again, it is clear that this kind of individualism is committed to (IC).

The third and fourth kinds of individualism are less straightforward. They acknowledge that the Twin Earth thought experiments establish a sense in which S and counterfactual S have different thoughts, and hence a sense in which an individual’s representational mental kinds are constitutively dependent on her relations to a wider environment. But they also maintain that there is a sense in which S and counterfactual S have the same thoughts, and hence a sense in which an individual’s representational mental kinds supervene locally on her individual perceptual discriminatory capacities. The first of these views is a two-factor theory according to which representational mental states and events have two, essentially connected components.¹³ The individualistic component of a thought—its narrow content—is a function that determines, relative to an environment, its anti-individualistic component—its broad, truth-conditional content. Thus the narrow content of S’s thought is the same in both the actual and the counterfactual scenarios, but the broad content of her thought in the actual scenario is different from the broad content of her thought in the counterfactual scenario because she is related to different objective properties in each. On this view, the narrow component of thought is not itself representational—only the broad component is; but the narrow component is responsible for representation, and is typically taken to be the fundamental

scientific notion. It is taken to be fundamental because it is thought that scientific kinds are individuated in terms of causal powers, and that causal powers depend on the intrinsic properties of an individual and not on an individual's relational properties. The two-factor theory, then, is committed to (IC) in virtue of its prioritising an individualistic notion of content—content individuated in terms of an individual's intrinsic causal powers, which supervene locally on her individual perceptual discriminatory capacities.¹⁴

The fourth individualist position also draws a distinction between broad and narrow content, but, unlike the two-factor theory just discussed, it maintains that all content is representational. On this view, the broad content of a subject's thought is determined, in line with anti-individualist considerations, in part by relations she bears to objective properties in her environment. The narrow content of a thought, on the other hand, is individuated by the epistemic possibilities it allows and excludes.¹⁵ The underlying thought here is that intrinsic duplicates are in the same epistemic position in the sense that they cannot distinguish between the relevant actual and counterfactual situations and that the narrow content of a thought encapsulates this fact. According to this version of individualism, S and counterfactual S have thoughts with different broad contents—one concerns silver whereas the other concerns twilver—but the thoughts have the same narrow content because a purely qualitative description of a situation in which silver needs to be polished to prevent it from tarnishing is identical to a purely qualitative

¹³ See Fodor (1980), (1987) and (1991), McGinn (1982) and Stich (1978). See also Stich (1983). Fodor moves away from the view in his (1994), but remains committed to (IC).

¹⁴ The view is typically set within a physicalist framework, and so the causal powers are understood as identical to or underwritten by physical causal mechanisms. The framework is not essential to the view, however, and my discussion is not thus restricted.

¹⁵ See Chalmers (2003).

description of a situation in which twilver needs to be polished to prevent it from tarnishing. Given the epistemic position of S and counterfactual S, both situations ‘verify’ their thoughts and hence they share a narrow content. Since narrow content thus construed is individuated by epistemic possibility—rather than by metaphysical possibility—it is clear that this version of individualism is also committed to (IC).

I take it, then, that just as the common core of anti-individualist theories is their commitment to (EC), the common core of individualist theories is their commitment to (IC).

4. Subjective Externalism

Over the previous two sections we have seen that anti-individualism is committed to an external, objective constraint on the correct attribution of thought content, which I have formulated as (EC), and that individualism is committed to an internal, subjective constraint on the correct attribution of thought content, which I have formulated as (IC). Individualism and anti-individualism are inconsistent theories about the individuation conditions of representational mental kinds, states and events. And yet both constraints, (EC) and (IC), have some prima facie plausibility. Certainly, large numbers of philosophers find (EC) compelling, as evidenced by the widespread acceptance of anti-individualism in the literature. The simple, underlying thought here is that it is unclear how an individual could have any representational mental states at all without having interacted with an environment that contains instances of at least some of the objective

properties that those mental states come to represent.¹⁶ But there is also significant support for (IC) in the literature, as evidenced by the variety of individualist positions discussed in section 2 above; and the simple, underlying thought here is that it is mysterious how objective properties that in no way impinge on an individual's perspective could be represented by her in thought. In what follows I will argue that the anti-individualist intuition that the correct attribution of thought content to an individual is essentially constrained by her non-representational relations to objective properties in her wider reality is consistent with the individualist intuition that objective properties in the environment to which a subject is perspectively blind cannot enter into the individuation conditions of her representational mental kinds. That is, I will argue that (EC) and (IC) are consistent. The reconciliation comes from understanding that the properties that serve in part to individuate one's mental states and events can be both objective properties in one's environment and properties that supervene locally on one's individual perceptual discriminatory capacities—that they need not be properties to which one is perspectively blind. The resulting view, which I call 'subjective externalism', is a form of anti-individualism that maintains that the correct attribution of thought content to an individual is sensitive to her individual perceptual discriminatory capacities.¹⁷

¹⁶ Anti-individualism does not require that all of an individual's concepts represent objective properties in her environment. The anti-individualist claim is simply that there could be no representation in thought without *some* relations to objective properties in an environment.

¹⁷ Note that the third and fourth of the individualist theories discussed above each admit one kind of thought content that satisfies (EC), and a different kind of thought content that satisfies (IC). Subjective externalism, in contrast, admits of one form of thought content that satisfies both.

There is an inconsistency in the neighbourhood, of course, since, as mentioned previously, (IC) is inconsistent with a particular kind of anti-individualism, namely essentialist externalism. But this does not mean that (IC) is inconsistent with anti-individualism *per se*, since essentialist externalism, as will become clear, is just one form of anti-individualism. Anti-individualism need not involve a commitment either to the claim that content is finer-grained than the individual perceptual discriminatory capacities of the subject, or to the claim that content is sensitive to objective properties to which the subject is perspectively blind. Anti-individualism is merely committed to the claim that many representational mental kinds, states and events are constitutively what they are partly by virtue of relations between the individual in those states and a wider reality.

At the heart of subjective externalism is the simple idea that objects and substances present themselves to us by way of their manifest properties and not, at least initially, by way of their essences or individual natures. To invoke Peacocke's (1993) metaphor, the identity of objects, or the haecceity of substances is 'bleached out'.¹⁸ Subjective externalism maintains that the individuation of a subject's representational mental kinds depends essentially on relations between her and objective properties in her environment that are manifest to her. The twin notions of a manifest property and perspectival blindness are opposite sides of the same coin, and can be inter-defined as follows. A property is manifest to a subject S at time t if and only if it is not a property to which she

¹⁸ According to subjective externalism, and in contrast to Peacocke, the identity of objects and the haecceity of substances need not be forever 'bleached out'. There is no in-principle reason to think there are properties to which human subjects are necessarily perspectively blind, although there may well be properties to which every human subject in fact remains perspectively blind.

is perspectively blind at time t . The notion of a manifest property can be characterized in its own terms as follows:

(M) A property P is manifest to a subject S at time t if and only if every property P' distinct from P , is such that S has the individual perceptual discriminatory capacity at t to distinguish some instances of P from some instances of P' .

Being manifest is a relation between a subject and a property at a time, and so being manifest is subject- and time-relative. A property which is not manifest to a subject at one point in time may be manifest to her at a later point in time, since her individual perceptual discriminatory capacities may change in relevant ways. And once a property becomes manifest to a subject, according to subjective externalism, it can then bear a constitutive, individuating relation to her representational mental kinds, states and events. Subjective externalism can be characterized as follows:

(SE) The correct attribution of thought content to an individual is essentially constrained by her non-representational relations to objective manifest properties in her wider reality.

Subjective externalism is clearly committed to (IC). According to subjective externalism, if the manifest properties of an individual's environment are held constant, her representational mental kinds will also remain constant, no matter how much the properties to which she is perspectively blind are altered. And since manifest properties

are defined in terms of individual perceptual discriminatory capacities, the local supervenience thesis that defines (IC) is upheld: a subject's mental states and events supervene locally on her individual perceptual discriminatory capacities. Subjective externalism is also committed to (EC). (SE) in fact differs from (EC) only by the introduction of the term 'manifest'. However, a more detailed account of the notion of a manifest property and of the ways in which properties can become manifest to a subject over time is needed in order to establish that manifest properties can be genuinely objective.

In what follows I identify three different elements that can play a role in making a property manifest to an individual: unaided perception; perception aided by the use of instruments; and theoretical knowledge. I will take each in turn, and then briefly clarify the relations between them. I frame the discussion in terms of three corresponding levels. At the first level, the properties that are manifest to a subject are the properties that the subject can discriminate by means of her individual perceptual systems, unaided by instruments or theory. The kinds of properties that a typical human subject can discriminate in this way are enormously diverse and, on the face of it, include qualitative properties, such as properties related to shape, colour, sound, texture and taste, artifactual properties, such as the properties of being a clock, a computer, or a steam engine, some dispositional properties, such as the properties of being fragile or being precariously balanced, certain psychological properties including emotional properties, propositional attitude properties and character traits, and some temporal properties, such as the property of being about to fall, or the property of having been freshly painted. The important feature of these properties is that instances of them do not have systematic manifest

doubles. Contrast the property of being silver: since something's being silver depends essentially on a structure that is hidden from view, the property of being silver does, for most individuals, have instances with systematic manifest doubles (instances of twilver), but the properties listed above plausibly do not.

There may be disagreement over particular categories or particular cases, but it is clear that the properties that are manifest to a subject at this fundamental perceptual level need not be restricted to secondary properties in Locke's sense¹⁹ or to the kinds of properties with which Russell thought we were acquainted through our sense-data.²⁰ Such a restriction would be unmotivated without an independent commitment to a foundationalist, infallibilist epistemology. And not only is there reason to question such a restrictive epistemology, but subjective externalism bears no such commitment.

The characterization of a manifest property given in (M) is consistent with a subject's being fallible with respect to a property which is manifest to her. For example, suppose Oscar has a naughty nephew with a practical joke kit containing stickers which, when put on glass, make it look as if the glass is cracked. The naughty nephew puts a joke sticker on the windscreen of Oscar's car one day when Oscar is out. When Oscar returns home he walks up the street with a clear view of his car windscreen. On approaching, he becomes more and more alarmed as he sees what he takes to be a crack in his windscreen. When he gets close enough, however, he realises his windscreen is not cracked and that he's been tricked by a simple sticker. He is relieved about his windscreen, and immediately turns his attention to finding his naughty nephew. The example illustrates that Oscar is fallible with respect to the property of being a cracked windscreen, but the

¹⁹ See Locke (1690), especially Book 2, chapter viii.

property is nonetheless manifest to him. We can see this by substituting Oscar and the property of being a cracked windscreen into (M) to yield (MO):

(MO) The property of being a cracked windscreen is manifest to Oscar at time t if and only if every property distinct from it (being a sticker, being a computer, etc.), is such that Oscar has the individual perceptual discriminatory capacity at t to distinguish some instances of being a cracked windscreen from some instances of every property distinct from it (being a sticker, being a computer, etc.).

The kind of fallibility that a subject has with respect to properties that are manifest to her is a kind that is rectifiable on the basis of her current individual perceptual discriminatory capacities. As Oscar discovered, look closely and the game is up. In contrast, the kind of fallibility that a subject has with respect to properties to which she is perspectively blind is systematic—that is, not rectifiable on the basis of her current individual perceptual discriminatory capacities. For example, there is nothing that our subject S can do on the basis of her current individual perceptual discriminatory capacities to discover that her bracelet is silver rather than twilver. She can acquire the capacity as an individual to distinguish silver from twilver; but this would necessarily involve a change in her current individual perceptual discriminatory capacities. Given that the notions of a manifest property and perspectival blindness are opposite sides of the same coin, systematic

²⁰ See for example Russell (1912).

manifest doubles occur, by definition, only when a subject is perspectively blind to the property in question.²¹

At the second level, the properties that are manifest to a subject are the properties she can distinguish on the basis of her individual perceptual systems but this time by means of instruments. A typical human subject does not have the individual perceptual discriminatory capacities to allow her to identify, unaided by the use of instruments, the property of weighing 75 grams, but she does have the individual perceptual discriminatory capacities to identify that property if she is allowed, still as an individual, to use a set of weighing scales. This capacity, of course, depends essentially on the capacities of her individual perceptual system, but is one step removed from the fundamental perceptual level by requiring instruments. Many instruments, such as weighing scales, rulers, thermometers, metronomes, telescopes and microscopes are relatively commonplace, and can with ease extend the range of objective properties that are manifest to the average human subject. This is in part because the properties identified by means of such instruments are precisifications of properties that are already manifest at the fundamental perceptual level. For example, at the fundamental perceptual level objects can be identified as differing in weight, and so units of measurement can be introduced together with an instrument that allows finer discriminations along what is at root the same scale.

²¹ It has been argued that the content of perceptual experience is open to Twin Earth thought experiments. See for example Chalmers (2004) and (2006). This might lead one to believe that there are no properties that are not subject to systematic manifest doubles, which would be problematic for subjective externalism. However, such Twin Earth thought experiments seem to depend on a representational theory of perceptual content and seem to assume essentialist externalism. As such, I leave them to one side here, although further discussion is warranted. In relation to this issue, see also Hellie (2014).

Increasingly complex scientific instruments are developed to identify properties that are increasingly removed from the properties that are manifest at the fundamental perceptual level. Examples include an ammeter, which measures electrical current, an electrostatic analyser, which measures the kinetic energy of charged particles, a magnetometer, which measures magnetic flux, an ohmmeter, which measures electrical resistance, particle accelerators, which indicate the nature of interactions between particles, and so on. Although such instruments are designed to identify properties which are relatively far removed from the properties which are manifest at the fundamental perceptual level, there is reason to regard them as extensions of the more commonplace instruments, and hence as, for the individuals that use them, extending the range of objective properties that are manifest in a similar way.

Again, the important feature of the properties that are manifest at this level is that instances of them do not have systematic manifest doubles. There may be disagreement over particular cases, but the existence of a class of properties which are manifest only by means of instruments need not be ruled out in principle by a prior commitment to a restrictive epistemology such as that evident in van Fraassen's constructive empiricism.²² According to van Fraassen, the use of instruments in science allows us to make predictions about the observable world, but does not allow us to perceive properties we could not otherwise perceive. For example, according to van Fraassen, a scientist does not see an abnormal blood cell when she looks through a microscope—what she sees is an image. The property of being an abnormal blood cell is therefore not, if van Fraassen is right, manifest to a subject when she looks through a microscope. But subjective

²² Van Fraassen (1980).

externalism is not committed to the restrictive kind of epistemology that underpins constructive empiricism. The characterization of a manifest property given in (M) is not only consistent with a subject's being fallible with respect to a property which is manifest to her, it is also consistent with the use of instruments improving a subject's discriminatory capacities rather than defining the limits for them. Nonetheless, it must be remembered that a property which can be made manifest by means of an instrument will only in fact be manifest to a subject who uses the relevant instrument herself. This ensures the preservation of the local supervenience of mental states and events on a subject's individual perceptual discriminatory capacities. There is a fundamental difference between people and instruments in this regard which means that it is not always possible to extend one's individual discriminatory capacities by using other people as instruments. Properties do not necessarily become manifest through testimony. This is because while the use of instruments *depends* upon one's own individual perceptual discriminatory capacities improving (since without the necessary improvement, one would be unable to use the instrument), the use of testimony does not depend upon one's own individual perceptual discriminatory capacities improving (although it may on occasion prompt a subsequent improvement). The idea that objective properties can be made manifest by the use of instruments is similar in spirit to Rouse's claim that the use of instruments in science allows parts of the world we otherwise could not observe to disclose themselves to us.²³ This provides an apt metaphor for the second level.

²³ Rouse (1987).

At the third level, the properties that are manifest to a subject are properties she cannot discriminate without bringing to bear her theoretical knowledge. It is important to note here that the three elements I have identified (unaided perception; perception aided by the use of instruments; and theoretical knowledge) need not be distinct in practice. In particular, the third element—theoretical knowledge—clearly plays a role in a subject's understanding of what she perceives by means of complex scientific instruments, and hence theoretical knowledge clearly plays a role at the second level described above. After all, theory must be brought to bear in order to interpret the data provided by the relevant instruments, and this will be increasingly necessary as the instruments make manifest properties that are increasingly distant from properties that are manifest at the fundamental level. For example, the ability to distinguish silver from twilver, or water from twater, lies firmly at the third level.²⁴ But theory may also play a role when a subject relies on a more commonplace instrument, and may even play a role in certain cases at the first level of unaided perception. Despite the fact that theoretical knowledge may be essentially involved in at least some cases at both the first and second levels, according to subjective externalism it is the level of non-theoretical, unaided perception that remains fundamental. This is because non-representational, causal relations between a subject and objective properties in her environment form the basis of all representation in thought. As such, all of a subject's discriminatory capacities that require the use of theoretical knowledge will depend asymmetrically on her individual, unaided perceptual

²⁴ In fact, terms such as 'twilver' and 'twater' are dummy predicates that fail to express natural kind concepts and fail to refer to natural kinds. In order to appreciate the point of the examples, we need to make-believe that such terms refer to natural kinds that are distinct from but superficially identical to natural kinds in our environment. For more on this see Sawyer (2015).

discriminatory capacities.²⁵ Indeed, it is this asymmetric dependence on the level of unaided perception that ensures that subjective externalism complies both with the externalist, objective constraint on the correct attribution of thought content captured by (EC), and with the internalist, subjective constraint on the correct attribution of thought content captured by (IC).

Subjective externalism is a version of what Davies calls ‘constitutive externalism’, since it is committed to the claim, ‘... that the fundamental philosophical account of what it is for an individual to have [a mental property] M does need to advert to the individual’s physical ... environment’ (Davies, 1998: 327). But the features of the environment that such an account needs to advert to are, according to subjective externalism, objective properties in the environment that are manifest to the individual concerned. And as her practical and theoretical knowledge increases over time, more of the world becomes manifest to her. Subjective externalism, then, is consistent with both (EC) and (IC).

Since subjective externalism is consistent with both (EC) and (IC), it is unsurprising to find that it agrees with essentialist externalism on some points and with individualism on others. As an example of the former, subjective externalism accepts that a traditional brain-in-a-vat will lack mentality because it lacks non-representational relations to objective properties in a wider reality; a traditional brain-in-a-vat fails to meet condition (EC). This accords with the intuition that a brain-in-a-vat does not have individual

²⁵ This mirrors the reconciliation of two anti-individualist claims that might be thought to be in tension: first, the claim that representation depends on non-representational relations to objective properties; and second, the claim that some representations do not represent objective properties. The reconciliation

perceptual discriminatory capacities.²⁶ As an example of the latter, subjective externalism accepts that individuals identical in their individual perceptual discriminatory capacities will have the same representational content even if they stand in non-representational relations to different wider realities. This is because of the commitment to (IC). One (perhaps welcome) consequence of this particular aspect of the view is that subjective externalism avoids having to deal with issues raised by cases of slow-switching.²⁷ Thus at the heart of subjective externalism lies the thought that although representation requires non-representational relations to objective properties in an environment, any such objective property is potentially manifest, and hence what is initially ‘external’ can be ‘internalized’, as we might put it.

5. Conclusion

In this paper I have tried to carve out a place for a theory of representational mental kinds which respects two apparently inconsistent constraints on the attribution of thought content. On the one hand we have an internal, subjective constraint, according to which the attribution of thought content to an individual is essentially constrained by her individual perceptual discriminatory capacities. This is the constraint to which individualistic theories of representational mental kinds are committed. On the other hand we have an external, objective constraint, according to which the attribution of thought content to an individual is essentially constrained by her non-representational relations to

depends on understanding that representations that do not represent objective properties depend asymmetrically on representations that do.

²⁶ A Putnam-style brain-in-a-vat adds further complexity and requires separate treatment. See Putnam (1981) for the example.

objective properties in her wider reality. This is the constraint to which anti-individualistic theories of representational mental kinds are committed. Subjective externalism acknowledges both constraints. This, then, is a view that reconciles the motivation behind individualist theories with the motivation behind anti-individualist theories. In the end, there may be reasons that favour essentialist externalism over subjective externalism. But subjective externalism does have the merit of reconciling a fundamental intuition that drives individualism with a fundamental intuition that drives anti-individualism. And this is a merit that should not, I think, be dismissed lightly.

References

- Burge, Tyler. 1979. Individualism and the mental. *Midwest Studies in Philosophy* 4: 73-121.
- . 1982. Other bodies. In A. Woodfield, ed., *Thought and Object: Essays on Intentionality*. Oxford: Oxford University Press.
- . 1986. Intellectual norms and foundations of mind. *Journal of Philosophy* 83: 697-720.
- . 1988. Individualism and self-knowledge. *Journal of Philosophy* 85: 649-63.
- . 2007. *Foundations of Mind*. Oxford: Oxford University Press.
- Chalmers, David. 2003. The nature of narrow content. *Philosophical Issues* 13: 46–66.
- . 2004. The representational character of experience. In Brian Leiter, ed., *The Future for Philosophy*. Oxford: Oxford University Press.

²⁷ For the original discussion of slow-switching see Burge (1998).

- . 2005. Perception and the fall from Eden. In T. Gendler and J. Hawthorne, eds, *Perceptual Experience*. Oxford: Oxford University Press.
- Davies, Martin. 1998. Externalism, architecturalism, and epistemic warrant. In C. Wright, B.C. Smith and C. Macdonald, eds, *Knowing Our Own Minds*. Oxford: Oxford University Press.
- Descartes, Rene. 1961. *Meditations on First Philosophy*, translated by John Cottingham. Cambridge: Cambridge University Press, 1996.
- Evans, Gareth. 1982. *The Varieties of Reference* J. McDowell , ed. Oxford: Oxford University Press.
- Fodor, Jerry. 1980. Methodological solipsism considered as a research strategy in cognitive psychology. *Behavioural and Brain Sciences* 3: 63-109.
- . 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA.: MIT Press.
- . 1991. A modal argument for narrow content. *Journal of Philosophy*, 88: 5-26.
- . 1994. *The Elm and the Expert*. Cambridge, MA.: MIT Press.
- Hawthorne, John. 2007. Direct reference and dancing qualia. In T. Alter and S. Walter, eds, *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Oxford: Oxford University Press.
- Hellie, Benj. 2005. Noise and perceptual indiscriminability. *Mind* 114: 481-508.
- . 2014. Love in the time of cholera. In B. Brogaard , ed., *Does Perception Have Content?* Oxford: Oxford University Press.
- Kripke, Saul. 1972. *Naming and Necessity*. Oxford: Basil Blackwell Ltd..

- Locke, John. 1690. *An Essay Concerning Human Understanding*. London: Thomas Bassett.
- McDowell, John. 1977. On the sense and reference of a proper name. *Mind* 86: 159-85.
- . 1984. *De re* senses. In Wright, C. , ed., *Frege: Tradition and Influence*. Oxford: Basil Blackwell.
- . 1986. Singular thought and the extent of inner space. In P. Pettit and J. McDowell, eds, *Subject, Thought and Context*. Oxford: Clarendon Press.
- McGinn, C. 1982. The structure of content. In Woodfield, A., ed., *Thought and Content*. Oxford: Oxford University Press.
- Peacocke, Christopher. 1993. *Sense and Content*. Oxford: Clarendon Press.
- Putnam, Hilary. 1973. Meaning and reference. *Journal of Philosophy* 70: 699-711.
- . 1981 *Reason, Truth and History*. Cambridge: Cambridge University Press.
- . 1996. Introduction. In Pessin, A. and Goldberg, S., eds, *The Twin Earth Chronicles: Twenty Years of Reflection on Hilary Putnam's 'The Meaning of "Meaning"'*. New York: M. E. Sharpe, Inc.
- Rouse, Joseph. *Knowledge and Power: Toward a Political Philosophy of Science*. Ithaca: Cornell University Press, 1987.
- Russell, Bertrand. 1912. *The Problems of Philosophy*. London: Williams and Norgate.
- Sawyer, Sarah. 2007. There is no viable notion of narrow content. In B. McLaughlin & J. Cohen, eds, *Contemporary Debates in Philosophy of Mind*. Oxford: Blackwell Publishers Ltd.

- . 2011. Internalism and externalism in mind. In J. Garvey , ed., *The Continuum Companion to the Philosophy of Mind*. London: Continuum International Publishing Group.
- . 2014. Contrastive self-knowledge. *Social Epistemology* 28: 139-152.
- . 2015. Contrastive self-knowledge and the McKinsey paradox. In S. Goldberg, ed., *Externalism, Self-Knowledge and Skepticism/* Cambridge: Cambridge University Press.
- . 2018. On the importance of concepts. *Proceedings of the Aristotelian Society* 118 (2): forthcoming.
- Segal, Gabriel. 2000. *A Slim Book about Narrow Content*. Cambridge, Mass.: MIT Press.
- Stich, Stephen. 1978. Autonomous psychology and the belief-desire thesis. *Monist* 6: 573-91.
- . 1983. *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge, Mass.: MIT Press.
- Van Fraassen, Bas. 1980. *The Scientific Image*. New York: Oxford University Press.