

What Good Are Our Intuitions?
Philosophical Analysis and Social Kinds

Sally Haslanger

MIT Department of Linguistics and Philosophy

1. Introduction¹

Across the humanities and social sciences it has become commonplace for scholars to argue that categories once assumed to be “natural” are in fact “social” or, in the familiar lingo, “socially constructed”. Two common examples of such categories are *race* and *gender*, but there many others. One interpretation of this claim is that although it is typically thought that what unifies the instances of such categories is some set of *natural* or *physical* properties, instead their unity rests on *social* features of the items in question. Social constructionists pursuing this strategy—and it is these social constructionists I will be focusing on in this paper—aim to “debunk” the ordinary assumption that the categories are natural, by revealing the more accurate social basis of the classification.² To avoid confusion, and to resist some of the associations with the term ‘social construction’, I will sometimes use the term ‘socially founded’ for the categories that this sort of constructionist reveals as social rather than natural.³

Let me emphasize: the idea in saying that a category is socially founded is not to say that social factors are responsible for our attending to the category in question (which may be true of wholly natural categories); nor is it to say that the things in the category are less than fully real (material things may be unified by social features and there is no reason to deny that social properties and relations are fully real). The point, roughly, is to shift our understanding of a category so we recognize the real basis for the unity of its members. As we will see, there are importantly different sorts of cases. But because the difference between a natural and a social category has significance both for what’s possible and for what we’re responsible, the constructionist’s general project, when successful, has important normative implications.

Amongst those who aim to analyze our ordinary racial classifications, social constructionists are often at odds with *error theorists* (sometimes called *eliminativists*) and *naturalists*. Error theorists maintain, in agreement with social constructionists, that the items taken to fall within the category in question do not meet the supposed natural or physical conditions for membership; the error theorist often goes farther to claim that the conditions are vacuous: nothing satisfies them, (sometimes even that nothing *could* satisfy them). They conclude, then, that such things are illusory and that talk purporting to refer to such categories is false or misguided. So, for example, an error theorist about race (Appiah 1996; Zack 1997) claims that there are no races, given what we mean by ‘race’. Of course it is then open to the error theorist to propose terminology for new categories—perhaps social categories with an extension close to what we thought was the extension of our original categories, such as “racial identities” (Appiah 1996)—whose conditions for membership are satisfied.

Present-day race *naturalists* agree with the eliminativists and constructivists that races are not what they were once thought to be—they are not, e.g., groups with a common racial essence that explains a broad range of psychological and moral features of the group’s members—but they disagree with both other views in maintaining that the human species can be divided on the basis of natural (biological, genetic, physical) features into a small set of groups that correspond *roughly* to the ordinary racial divisions (Rosenberg et al 2002; Mountain and Risch 2004). This in itself would not be particularly interesting, however, if the natural basis for the grouping was biologically “real” but of no real significance for explanation or prediction. Full-blooded race naturalists, however, maintain that there is a biologically significant classification that somehow captures our current racial divisions. Moreover, they argue, recognizing this fact is socially and politically important for the purposes of achieving racial justice, e.g., by enabling us to address racially divergent medical needs (Risch et al 2002; cf. Lee et al 2001).⁴

In this essay I shall focus on the debate between the constructionist and error theorist; I take up the disagreement between the constructionist and naturalist elsewhere (Haslanger

2006). One way of capturing the difference between the social constructionist and the error theorist is to see them as disagreeing about the content of the relevant concepts. According to the error theorist, there is reason to take our ordinary belief that the category has a natural basis to set a constraint on what could count as an adequate analysis of the concept: a successful analysis must be in terms of natural properties and relations (or involve them at least in the way required for the concept to count as expressing a natural property). In contrast, according to the social constructionist, we may employ the concept successfully even though we have a false belief about what sort of property it expresses or sort of set it determines. For example, a social constructionist about race (such as myself) will claim that there are races and that races are social categories, i.e., that race is socially founded, even though it is commonly assumed that races are natural categories. It is an important part of the social constructionist picture that, to put it simply, our meanings are not transparent to us: often ideology interferes with an understanding of the true workings of our conceptual framework and our language. More specifically, ideology (among other things) interferes with our understanding of our classificatory practices, suggesting to us that we are finding in nature divisions that we have played an important role in creating.

The Concept of Race

To gain a vivid sense of the controversy, it may be helpful to consider briefly two different analyses of *race*, one constructionist, the other defended by the error theorist. These are just two examples of many that are discussed in the literature.

Anthony Appiah is perhaps the most well-known error theorist about race. On his view the concept of race is the core notion in the folk theory of racialism:

...there are heritable characteristics, possessed by members of our species, that allow us to divide them into a small set of races, in such a way that all the members of these races share certain traits and tendencies with each other that they do not share with members of any other race. These traits and tendencies characteristic of a race constitute, on the racialist view, a sort of racial essence; and it is part of the content of racialism that the

essential heritable characteristics...account for more than the visible morphological characteristics—skin color, hair type, facial features—on the basis of which we make our informal classifications. (Appiah 1993, 5)

Races, then, are groups with a common inherited racial essence. The implications of this for our purposes are straightforward: there are no such racial essences, so there are no races. Appiah argues, however, that there is a neighboring notion—that of *racial identity*—that does not presuppose racial essences and can be accurately attributed to people. Having a racial identity is a matter of identifying with a label (such as “White” or “Black”) that has been historically associated with a racial essence (Appiah 1996, 81-2).

In contrast, I have argued for a constructionist account of race that parallels an account of gender (Haslanger 2000). On my view (to simplify quite a bit) races are racialized groups, and:

A group is *racialized* (in a context) if and only if its members are socially positioned as subordinate or privileged along some dimension—economic, political, legal, social, etc.—(in that context), and the group is “marked” as a target for this treatment by observed or imagined bodily features presumed to be evidence of ancestral links to a certain geographical region.

On this view, being White (in a context) is a matter of being seen as conforming to a meaningful bodily schema associated with European ancestry—such schemas I call “color”—and being treated (in that context) as positioned in a social hierarchy appropriate for persons of that “color”. In the contemporary United States, being marked as “white” brings with it a broad range of social privileges, at least for the most part.⁵ However, because racial hierarchies interact with other social hierarchies—gender, class, sexuality, culture, religion, nationality—the concrete impact of being White varies depending on other aspects of one social position. For example, a straight young White man and an elderly White lesbian will both reap privileges by virtue of their Whiteness, but the kinds of privileges they enjoy may differ considerably.

Three questions naturally arise at this point: First, it is clear that the analysis of race I offer does not capture what people consciously *have in mind* when they use the term 'race'. The account is surprising, and for many, highly counter-intuitive. (Although I myself doubt that Appiah's account captures better what people consciously *have in mind* when they talk of races, it is at least familiar, and has some intuitive plausibility if we are looking for how people generally have thought of races.) Note that this counter-intuitiveness will always be a feature of social constructionist analyses because (debunking) social constructionists aim to reveal that the concepts we employ are not exactly what we think they are. But if the adequacy of a philosophical analysis is a matter of the degree to which it captures and organizes our intuitions, and if constructionist analyses are always counter-intuitive, then it would seem that philosophers would never have reason to consider social constructionist projects acceptable. However, this seems too fast. Surely philosophers cannot simply rule out constructionist analyses from the start.

Second, does the social constructionist approach make sense? Are there considerations developed in the context of philosophy of language (or related areas) that would prevent one from pursuing a constructionist *analysis* of race or gender and force us to adopt an error theory? I will argue that, in fact, there are considerations in contemporary philosophy of language that not only permit, but in some respects *favor* a constructionist account.

Third, what difference does it make? Does it really matter whether we say, for example, that there are no races but there are racial identities, rather than there are races but they are social rather than natural? Are there cases where an error theory would be mistaken but a social constructionist account would be warranted? By what criteria do we decide, and is it worth worrying about?

The arguments that follow focus mainly on the first two questions; however, drawing on this discussion I will return to the question why it might matter whether we adopt a constructionist or error strategy towards the end of the essay. The issues are complex. My own view is that which approach is better will depend on the case at issue, and the betterness will

depend both on semantic, pragmatic, and political considerations. Moreover, pragmatic and political factors will vary with context. Before we proceed, however, it is worth pointing out that one potential advantage of a constructionist account is that it does not simply deny the existence of the allegedly natural category and substitute another (possibly social category) in its place, but it also—at least in the best cases—provides a diagnosis of our role in bringing about the effects that appear to us (mistakenly) as natural together with an explanation of the illusion. In such cases, the self-deception involved when we mean something, and yet mask that meaning to ourselves, is laid bare. Such unmasking can be an important step in motivating social change.

2. *Kinds of analysis*

The project of “conceptual analysis,” in philosophy takes many forms, partly depending on the particular concept in question, and partly depending on what methodological assumptions the philosopher brings to the issue. There are at least three common ways to answer “What is X?” questions: *conceptual*, *descriptive*, and *ameliorative*.⁶

For example, consider the question, “What is knowledge?” Following a *conceptual approach*, or what we might more revealingly call an *internalist approach*, one is asking: What is *our* concept of knowledge? and looks to *a priori* methods such as introspection for an answer.⁷ Taking into account intuitions about cases and principles, one hopes eventually to reach a reflective equilibrium. On a *descriptive* approach, one is concerned with what objective types (if any) our epistemic vocabulary tracks.⁸ The task is to develop potentially more accurate concepts through careful consideration of the phenomena, usually relying on empirical or quasi-empirical methods. Scientific essentialists and naturalizers, more generally, start by identifying paradigm cases—these usually function to fix the referent of the term—and then draw on empirical (or quasi-empirical) research to explicate the relevant kind or type to which the paradigms belong. Do paradigms project an objective type, and if so, what type? Familiar descriptive approaches in philosophy of mind and epistemology draw on cognitive science.

Ameliorative projects, in contrast, begin by asking: What is the point of having the concept in question, e.g., why do we have a concept of knowledge or a concept of belief? What concept (if any) would do the work best? In the limit case a theoretical concept is introduced by stipulating the meaning of a new term, and its content is determined entirely by the role it plays in the theory. If we allow that our everyday vocabularies serve both cognitive and practical purposes that might be well-served by our theorizing, then those pursuing an ameliorative approach might reasonably represent themselves as providing an account of our concept—or perhaps the concept we are reaching for—by enhancing our conceptual resources to serve our (critically examined) purposes (see Anderson 1995). Conceptual, descriptive, and ameliorative projects cannot, of course, be kept entirely distinct, but they have different subject matters and different goals.

Given the different projects of analysis and different subject matters for “analysis”, it is not surprising that philosophers who may appear to be asking the same question are in fact talking past each other. For example, where one philosopher might assume that an adequate analysis must capture our ordinary intuitions, another may take for granted that a priori reflection is likely to be systematically misleading when we are trying to understand the social domain. In fact, recent work on race provides an excellent example of the diversity of approaches. Some authors are engaged in a conceptual project, attempting to explicate our ordinary understanding of race (Appiah 1996, Zack 1997, Hardimon 2003, Mallon 2004); others are attempting to determine what, if any, natural kind we are referring to by our racial terms (Appiah 1996, Kitcher 1999, Andreason 2000, Zack 2002, Glasgow 2003); others have pursued genealogy (Omi and Winant 1994); still others are invested in what I call ameliorative projects, raising normative questions about how we should understand race, not only how we currently do (Gooding-Williams 1998, Alcoff 2000).

What should we make of these different projects? Should we simply allow that different inquirers are interested in different questions, and nothing can be said to resolve the question what race *really* is or what we mean by “race”? Although I would not argue that there is one

thing that race *really is* or one thing that “we” mean by ‘race’, we might hope that through reflection and discussion we could come to the point where (a) the concept we take ourselves to be employing, (b) the concept that best captures the type we are concerned with, and (c) the type we ought to be concerned with coincide. In such cases the conceptual, descriptive, and ameliorative projects yield the same concept. It is a mistake, then, for those engaged in conceptual analysis to dismiss other forms of analysis, with the thought that only the conceptual project can discover “our” concept (See Mallon 2004, Hardimon 2003). For example, if we discover that we are tracking something that is worthwhile to track in using our racial vocabulary, then even if this is not what we originally “had in mind,” it still may be what we have been and should continue to be talking about.⁹ But how should we proceed?

3. Manifest and operative concepts

One of the functions of concepts is to enable us to draw distinctions between things. Sometimes the activity of distinguishing things—separating them into groups—comes first, and we develop a concept of what we’ve distinguished later; sometimes the concept comes first, and we divide things according to it. For example, I might find myself asking my daughter to turn down her music on a regular basis without thinking that there is any pattern in my requests, only to find through conversation with her that I always ask her to turn it down when she is listening to a particular artist; or I might come to judge that a particular artist’s lyrics are typically offensive and ask her to turn down the music once I discern that that artist is playing. In practice, our activity of grouping things, even when we have a concept in mind, does not involve explicitly applying the concept to each case, i.e., making sure that each object meets the conditions for applying the concept. We typically rely on empirical assumptions linking easily accessible criteria with the conditions for membership. In the grocery, I pick up what look and feel like potatoes, without testing them genetically. Once I learn that the store sells genetically modified potatoes, I may want some further assurance of the genetic make-up of the ones I

purchase; but even then, I will rely on a sticker or label rather than applying the genetic criteria myself.

Everyday life requires a steady activity of drawing distinctions, an activity which combines both the use of concepts as guides and a rough-and-ready responsiveness to things. In reflecting back on our activity, there are a number of options for describing this sort of give and take. Consider again my requests to Zina (my daughter) that she lower the volume of her music. Suppose I don't want to listen to music with misogynistic lyrics. I have a concept of misogynistic lyrics and I also have a rough and ready responsiveness to what she is listening to. When Zina complains about my interventions into her listening, I may come to find that my responses are not tracking misogynistic lyrics after all, even though that's the concept I was attempting to use to guide my interventions. Let's call the concept I thought I was guided by and saw myself as attempting to apply, the *manifest concept*. I find, in other words, that my manifest concept is not in accord with my practice of determining when she has to lower the volume of her music.

There are several ways to resolve the awkward position of having my self-representation, or my intentions, out of line with my practice.

(i) I can be more careful about my interventions so that I only make my request when the lyrics really are misogynist. This would be to change the occasions of my intervention to bring my responsiveness into accord with my manifest concept.

(ii) I might instead find that a different concept conforms to my pattern of interventions. I could find that my responses are prompted by, say, sexually explicit lyrics not misogynistic ones. Let's call the concept that best captures the distinction that I in practice draw, the *operative concept*. In such a case, I allow the operative concept to have priority over the (original) manifest concept in guiding my behavior; in doing so the operative concept becomes manifest (and hopefully, is now consistent with my practice).

(iii) Rather than replacing the original manifest concept with a new operative concept, I modify my understanding of the manifest concept in light of the new cases that have emerged

in the practice. So rather than being newly guided by the concept of sexually explicit lyrics, I change what I understand misogynistic lyrics to be. Let's call the *target concept* the concept that, all things considered (my legitimate purposes, the facts, etc.), I should be employing. In the ideal case, I adjust my practice and my self-understanding to conform to the target concept.

So far I've distinguished the manifest, the operative, and the target concept. The manifest concept is the concept I take myself to be applying or attempting to apply in the cases in question. The operative concept is the concept that best captures the distinction as I draw it in practice.¹⁰ And the target concept is the concept I should, ideally, be employing. As illustrated above, the operative concept may not correspond with my understanding of what distinction I'm tracking. This is not to say, however, that the manifest and operative concepts always, or even typically, come apart. Typically, my practice will track the objective type that my manifest concept determines; in other words, my manifest concept and my operative concept coincide.¹¹ In the best cases, all three (my manifest, operative, and target concepts) will coincide.

The example of my responses to Zina's music locates the issue in the realm of individual consistency: how can I bring my practice in line with my intentions? To see that the phenomenon has broader scope, it may be helpful to consider an example that draws on more collective meanings. Consider the term 'parent'. It is common, at least in the United States, to address primary school memos to 'Parents,' to hold a "Parent Night" or "Parent Breakfast" at certain points during the school year, to have "Parent-Teacher Conferences" to discuss student progress, etc. However, in practice the term 'parent' in these contexts is meant to include the primary caregivers of the student, whether they be biological parents, step-parents, legal guardians, grandparents, aunts, uncles, older siblings, informal substitute parents, etc. However, it is also clear to everyone that those on the list just given are not the student's parents. So, for example, Tara's grandmother Denise (with whom she lives) counts as Tara's parent in all relevant school contexts, but is also known to be her grandmother and not her mother, and so not her parent. Given the distinction between manifest and operative concepts,

it would seem that there are two different concepts of 'parent' here: parent as *immediate progenitor*, and parent as *primary caregiver*. Tara's grandmother satisfies the operative concept of *parent* but not the manifest one.

One might resist the idea that the manifest concept of *parent* is of biological mother or father, however, my own experience as an adoptive mother has convinced me that at least in many contexts the dominant understanding of 'parent' frames it as a biological notion. For example, if I were Zina's biological parent, I don't think I would ever be asked (by people who know us) "Do you know Zina's parents?". If one is uncomfortable with the assumption that the manifest concept of *parent* is biological, then we need only take the case to be describing a possible world in which the manifest concept of *parent* is more narrowly biological, possibly a world much like the US in an earlier era before adoption was common or legally institutionalized.

As in the earlier example of misogynist lyrics, there are three different responses to the gap between idea and practice in our use of the term 'parent' :

(i) Bring our practice in line with the manifest concept: insist that one must be an immediate progenitor of a student to participate in Parent Nights, Parent-Teacher Conferences, etc. (This option seems clearly misguided—not necessarily as a semantic matter, but as a social/political matter.)

(ii) Find a new manifest concept that better captures our practice: correct the memos so they are addressed to "Primary Caregivers".

(iii) Modify our understanding of the manifest concept, in this case, 'parent', to accord with our practice. This would involve a transition in our understanding from parent as a biological category to parent as a social category.

This example is intended to show that the distinction between manifest and operative concepts is one that concerns public meanings as much as individual beliefs and intentions (see also Haslanger 2005). If we ask: What is the concept of 'parent'? we have at least two places to look for an answer: the concept that speakers generally associate with the term, and the concept

that captures how the term works in practice. Although so far I've focused on the relatively transparent example of 'parent', of course there are many philosophically rich and more surprising examples available. Feminist and race theorists have been urging for some time that the proper target of analysis is not (or not simply) what we have in mind, but the social matrix where our concepts do their work. For example, Catherine MacKinnon says, "[The verb 'to be' in feminist theory] is a very empirical 'is'. Men define women as sexual beings; feminism comprehends that femininity 'is' sexual. Men see rape as intercourse; feminists say much intercourse 'is' rape." (MacKinnon 1987, 59). Charles Mills argues that the Enlightenment social contract is a Racial Contract (Mills 1997), and that an adequate analysis of personhood reveals that "all persons are equal, but only white males are persons." (Mills 1998, 70) Such analyses purport to show that our manifest understandings of crucial political notions are masking how the concepts in question actually operate. (See also Mills 1998, 139-166.)

4. Concepts, Conceptions and the like

Can we understand the manifest/operative distinction in terms of a more familiar distinction between concept and conception? It is not an unusual circumstance in philosophy to find that "the concept" we take ourselves to be analyzing is not "the concept" that the students seem to employ in their day to day practice. Undergraduates are competent users of terms such as 'knowledge,' 'justice,' and 'object,' and yet are surprised and resistant when they learn philosophical theories of knowledge, justice, and objects. One might argue that the philosophical theories are all false and the students are correct to reject them. But this is often not plausible. More plausible is that ordinary usage of a term doesn't require that one has thought carefully enough about the issues to develop consistent accounts of central concepts in one's repertoire. Cases such as these (and in general, accounts of language acquisition) support the idea that we need to distinguish various ways that individuals might be related to the concept, say, of knowledge or justice. For example, James Higginbotham distinguishes (Higginbotham 1998, 149-50):

i) possession

- merely possessing a word, and so being able to use it with its meaning;
- merely possessing a concept, so being able to deploy it, without having an accurate or full conception of it.

ii) tacit conception

- knowing the meaning of the word;
- having a full conception of a concept.

iii) explicit understanding

- having an adequate conscious view of the word's meaning;
- having an adequate conscious view of the nature of the concept.

There may be additional relevant distinctions, but these go some way to making room for competent use without full explicit understanding. Over the course of repeated use of a notion we develop conceptions of what we're talking about, but we might be misguided in various ways. Philosophical inquiry helps us develop more detailed, explicit, and adequate conceptions of our concepts.

But this doesn't seem to capture what's at issue in the cases we've been considering. How would we map the distinction between manifest and operative concepts on some pair of the distinctions between possession, tacit conception, explicit understanding? For example, consider Brenda, Tara's teacher. She has a fully conscious and explicit understanding of the concept *parent*; she also enacts a practice that is in some ways at odds with it. And she probably doesn't have a full or explicit conception of the rule that she employs in practice. But the distinction between manifest and operative concepts is not simply a distinction along the continuum of implicit-explicit, or uninterpreted-adequately interpreted. As we saw before, the manifest and operative concepts are at odds; they are, in a sense, competing with each other within the space of practical reason. For example, given the confidentiality laws in place, Brenda may find herself uncertain how or what to communicate with someone she knows to be a primary caregiver of a student, but who is not legally recognized as such. For all intents and

purposes the adult in question is the child's parent and no other parent is available, but legally speaking the adult is not the child's parent. In effect, fully developing and making explicit the operative concept does not necessarily yield the manifest concept. And as we saw above, this does not mean that we should reject the operative concept as a "misguided conception" in favor of the manifest.

5. Is *parent* socially founded?

Let us return to consider further the gap between the manifest concept of *parent* and the operative concept. A social constructionist in this case will plausibly claim that the category of *parent* is "socially founded". This means both that our manifest concept of *parent* (understood in biological terms) does not accord with our practices involving the notion of *parent* (which extends beyond the biological), and also that we would do well to modify our understanding of *parent* to involve a social dimension.

It appears that the constructionist could pursue more than one strategy for making the modification. One would be to simply replace the manifest concept of *parent* with the operative (e.g., *primary caregiver*), and appropriate the terminology of 'parent'. This would be to adjust, in a brute way, our understandings to conform to our practice. (I'll call this the descriptivist strategy.¹²) A second strategy would be to reflect on content of the manifest concept and the practice to come up with a concept of *parent* that best suits our needs and legitimate purposes. (Let's call this the ameliorative strategy.) A third strategy might be to argue for an ambiguity in the term, with one meaning tracking a social kind; the question then is whether a new term should be introduced or whether there are other ways of resolving the ambiguity.

Those favoring a conceptualist (or internalist) analysis, as I've described it, typically argue that neither of these constructivist approaches—the descriptivist or the ameliorative—is acceptable because both amount to changing the subject. *Our* notion of parent (or *our* notion of race) is of a biological category, and any modification that disrupts that assumption replaces *our* concept with a different one. In other words, 'parent' just means *immediate progenitor*, and if

we start using it to mean *primary caregiver* (or some more philosophically refined notion) then we have changed the meaning.¹³ So, social constructionists are wrong to say that *parent* is socially founded; what they are really saying is that a different concept, e.g., *primary caregiver*, is socially founded, which is obvious and not worth pointing out.¹⁴ Moreover, the social founding of *primary caregiver* poses no challenge to our assumptions concerning the concept of *parent*. Similarly a conceptualist concerned with race maintains that *our* concept of race is of a biological category, and nothing satisfies the biological conditions the concept requires. So the best we can do to capture the phenomena is to deny that there are races and invent (or appropriate) a new concept—such as *racial identity*—for the type tracked by our practices.¹⁵

I am willing to grant that each of these strategies could be reasonable in some context and we cannot decide on the basis of the simplified descriptions I've offered which is the best overall. However, the conceptualist sometimes maintains that his approach is the *only* reasonable approach and that the descriptive and ameliorative approaches cannot capture the meaning of our terms; the best they can do is propose new meanings.¹⁶ Is there reason to think that a descriptivist or ameliorative approach is simply misguided and that social constructionists aren't really doing philosophical analysis of the terms in question at all?

In the case of *parent*, at least, I think it is fairly clear that the concept has evolved and continues to evolve in response to the changing circumstances of family life; significantly for our purposes, it has changed from functioning as a natural category to functioning—at least in some settings—as a social category. (In some contexts the manifest concept of parent seems to allow that step-parents and adoptive parents are fully included; and increasingly there are contexts where it is no longer surprising if a child has two parents of the same sex.) How do we take this evolution into account in doing philosophical analysis? What does this “evolution” involve? By granting that the concept has evolved, are we conceding to the error theorist that we are analyzing a new concept, not the concept of parent? If we adjust the manifest concept so that a biological relation to the child is no longer necessary, isn't this “changing the subject” in the very sense the conceptualist is worried about?

There are at least two ways of thinking of what might be involved when a concept evolves in response to social context: on one hand it may be that the term 'parent' expresses a different concept than it once did. The change is a change in our language. On the other hand it may be that the concept of parent remains the same, but what we take to be the shape and content of that concept changes. Perhaps we once took it to be an essential feature of parents that their children were biologically related to them, but we have come to regard this as just an empirical generalization based on a limited survey of cases that does not hold necessarily. This is a change in our conceptual knowledge. In the next two sections I will consider each of these interpretations of "conceptual change." I will argue, first, that the constructionist is changing the subject, or changing our language; rather, the constructionist is revealing that our linguistic practices have changed in ways that we may not have noticed. Second, I will argue that although the constructionist suggests that we come to a new understanding of our concepts, this does not require replacing our old concept with a new one, but to understand our original concept better. I do not commit myself to one or another account of conceptual change here; I also want to leave it open that concepts change like other ordinary things, i.e., by altering.

6. *Semantic Externalism*

According to the first explanation of conceptual change, the change is in what concept our terms (such as 'race' or 'parent') express. This is plausibly understood as a semantic shift. It is not within the scope of this paper to take a stand in debates over meaning, e.g., whether the *meaning* of the term is a concept to whose content we have privileged access, or the term's extension, or a function from worlds to sets of objects, etc. The point I want to make is quite simple and should be familiar: whatever it is that determines the extension of our social kind terms, it isn't something to which we have privileged access through introspection. If the extension of the term changes over time, it is legitimate to postulate a change in what determines the extension. Those who are familiar with an externalist approach to language and mind will find little new in this section beyond the claim that externalist insights should be

applied to our thought and language about the social as well as the natural (so take yourself to be given permission to skip ahead). However the implications of externalism are much less commonly recognized in social and political philosophy, so I'll provide a quick summary here.

I've suggested that the error theorist typically invests in a conceptualist approach to analysis that emphasizes a priori reflection and ideas that are relatively accessible to introspection. I've also suggested that it is plausible to see this as an investigation of the manifest concept.¹⁷ In undertaking conceptual analysis of, say, *Fness*, it is typically assumed that it is enough to ask competent users of English under what conditions someone *is F*, without making any special effort to consult those whose daily lives are affected by the concept or who use the concept in practice. After all, if competent speakers know the meaning of their terms, then all that's needed is linguistic competence to analyze a term. A sophisticated internalist might want to allow that if one is sensitive to the possibility that in any actual circumstance there are competing meanings (often quite explicit) that structure alternative practices, then one could and should consider a broad range of speakers that are differently situated with respect to the phenomenon.

However, this approach to understanding race, gender, and other social kinds is not plausible if one takes into account arguments in philosophy of language over the past thirty years that call into question the assumption that competent users of a term have full knowledge of what the term means, i.e., that what's "in our heads" determines a term's referent. This assumption was already questioned once we considered the distinction between concept and conception (see section 4), but is further challenged by the tradition of semantic externalism. Externalists maintain that the content of what we think and mean is determined not simply by intrinsic facts about us but at least in part by facts about our environment. Remember: Sally and Twinsally both use the term 'water', but Sally means H₂O and Twinsally means XYZ (Putnam 1975b). Sally thinks she has arthritis in her thigh, and is wrong because 'arthritis' in her community is an ailment of the joints; Twinsally thinks she has arthritis in her thigh and is

right because ‘arthritis’ in her community is an ailment that is not confined to the joints (Burge 1979).

Most commonly, descriptive analyses—and the externalist picture guiding them—have been employed to provide a *naturalistic* accounts of knowledge, mind, etc; these seek to discover the *natural* (as contrasted with *social*) kind within which the selected paradigms fall. But it is possible to pursue a descriptive approach within a social domain as long as one allows that there are social kinds or types.¹⁸ In fact I’ve chosen to speak of ‘descriptive’ approaches rather than ‘naturalistic’ approaches for just this reason. Descriptive analyses of social terms such as ‘democracy’ and ‘genocide’, or ethical terms such as ‘responsibility’ and ‘autonomy’ are methodologically parallel to more familiar naturalizing projects in epistemology and philosophy of mind.

Of course, an externalist analysis of social terms cannot be done in a mechanical way and may require sophisticated social theory both to select the paradigms and analyze their commonality; in short, the investigation of social kinds will need to draw on empirical social/historical inquiry, not just natural science. Moreover, it is easily possible that the resulting analysis of the type is highly surprising. For example, it was not intuitively obvious that water is H₂O or that gold is an element with atomic number 79. It took sophisticated natural science to determine what the terms ‘water’ and ‘gold’ mean. Likewise it may take sophisticated social theory to determine what ‘parent’ or ‘Black’ means. In a descriptive project, intuitions about the conditions for applying the concept should be considered secondary to what the cases in fact have in common: as we learn more about the paradigms, we learn more about our concepts.

Externalism initially appeared in two forms, supported by the sorts of examples (‘water’, ‘arthritis’) just recited:

1. *Natural kind externalism* (Putnam 1973; Putnam 1975b; Kripke 1980): natural kind terms/concepts pick out a natural kind, whether or not we can state the essence of the kind, by virtue of the fact that their meaning is determined by ostension of a paradigm (or other means

of reference fixing) together with an implicit extension to “things of the same kind” as the paradigm.

2. *Social externalism* (Putnam 1975b; Burge 1979; Burge 1986): the meaning of a term/content of a concept used by a speaker is determined at least in part by the standard linguistic usage in his or her community.

It then became clear that externalist phenomena are not confined to natural kind terms (properly speaking) but occur quite broadly. For example, in the history of logic and math, inquiry can seem to converge on an idea or concept that we seemed to have in mind all along, even though no one, even the best minds, could have explicated it. (Leibniz’s early efforts to define the limit of a series is an example.) In such cases it is plausible to maintain that certain experts were “grasping a definite sense, whilst also failing to grasp it ‘sharply’.” (Peacocke 1998, 50)

Recognizing the possibility of reaching for a concept that is not quite within grasp provides us with a way to think about the ameliorative approach to analysis sketched above. In such cases we have perhaps a partial or vague understanding of the manifest concept and the operative concept picks out a relatively heterogeneous set, but nonetheless we can say that there is something we mean, a objective type we are approaching. I will use the term “target concept” for the concept that is plausibly what we are getting at, even if we poorly understand it; the target concept is the object of ameliorative analysis. Although Fregeans are apt to capture this by invoking objective senses that the inquirers “grasp”, an ontology of sparse objective properties will also do the work.

The upshot of this is that the basic strategy of natural kind externalism need not be confined to natural kinds (where it is assumed that things of the same natural kind share an essence). Externalism is an option whenever there are relatively objective types. The notion of objective type needed is not too mysterious: a set of objects is more an objective type by virtue of the degree of unity amongst its members beyond a random or gerrymandered set. Objectivity is only to be found in the natural world. There are objective types in every realm: social,

psychological, political, mathematical, artistic, etc.¹⁹ We might account for unity in various ways (Lewis 1983), but a familiar way I'll assume for current purposes is in terms of degrees of similarity; the similarity in question need not be a matter of intrinsic similarity, i.e., things can be similar by virtue of the relations (perhaps to us) they stand in. Roughly,

objective type externalism: terms/concepts pick out an objective type, whether or not we can state conditions for membership in the type, by virtue of the fact that their meaning is determined by ostension of paradigms (or other means of reference fixing) together with an implicit extension to things of the same type as the paradigms.

Sets of paradigms will typically fall within more than one type. To handle this, one may further specify the kind of type (type of liquid, type of artwork), or may (in the default?) count the common type with the highest degree of objectivity. For the purposes of capturing the operative concept, it is promising (as suggested before) to take the relevant type to be the one that we rely on in our best theory of the social/linguistic practice.

Descriptive projects adopt an externalist approach to content, i.e., they set out to determine the (an?) objective type, if any, into which the paradigms of a particular concept fall. Social constructionists can rely on externalist accounts of meaning to argue that their disclosure of an operative or a target concept is not *changing the subject*, but better reveals what we mean. By reflecting broadly on how we use the term 'parent', we find that the cases, either as they stand or adjusted through ameliorative analysis, project onto an objective social, not natural, type. So although we tend to assume we are expressing the concept of *immediate progenitor* by the term 'parent' in fact we are expressing the concept of *primary caregiver* (or some such); the constructionist shows us that our assumptions about what we mean are false, given our practice. This is not to propose a new meaning, but to reveal an existing one.

If one assumes with the conceptualist that the task of philosophical inquiry is simply to explicate through introspection what we think we mean in using a term, then almost any externalist inquiry will seem "revisionary". But the conceptualist approach to analysis is wedded to assumptions about mind and language that are certainly contested, if not outmoded.

We should also ask ourselves why, given the systematic use of terms such as ‘parent’ to track a social category, do we persist in thinking that the term picks out a natural category? Might ideology be playing a role in masking how we organize our social lives? In any case, there is no reason to reject out of hand the constructivist’s claim that a term whose manifest concept is of a natural kind may be better understood in terms of the operative social concept. The proposed analysis may be surprising; and it may even be that the term has come, through practice, to express a different concept than it used to, i.e., the manifest may not have caught up with the operative. But the constructionist is not causing this, or even promoting it such a change, but is rather revealing it.

7. *Meaning holism*

We considered before two ways that we might interpret the idea that a concept “evolves” with social practices. On one interpretation, the point is that a term such as ‘parent’ expresses, say, the concept of *immediate progenitor* at one time, but given changes in how the community organizes family life, comes to express a different concept, e.g., *primary caregiver*, at a later time. I’ve argued that if the constructionist describes this shift by saying that the term ‘parent’ is socially founded. This is not to invoke or propose a new meaning, but rather, drawing on externalist insights, to reveal an existing meaning that might well be obscured.

However, the other interpretation we considered was that the concept in question “evolves”, not in the sense that the term changes in what concept it expresses, but rather, there is a change in our understanding of the concept. For example, empirical investigation might reveal that a generalization we took to be analytically entailed by the concept is in fact only contingent, or even false. The suggestion here is not that the concept itself changes (though it might be useful to spell it out that way), but rather that our understanding of it does.

Nothing I’m saying here is news; the claim that there is, at best, a blurry line between what’s true by virtue of fact and what’s true by virtue of meaning is one theme in the arguments against the analytic/synthetic distinction. As in the case of externalism, however, the focus of

discussion has typically been on cases in natural science and the development of natural kind concepts. For example, in his essay, “The Analytic and the Synthetic,” Putnam contrasts examples in math and science which, he argues, are not happily classified as either analytic or synthetic, with the standard example of “a bachelor is an unmarried man” which is one of the few claims that he thinks should count as analytic (Putnam 1975a). He says:

In the case of a law cluster term such as ‘energy’, any one law, even a law that was felt to be definitional or stipulative in character, can be abandoned, and we feel that the identity of the concept has, in a certain respect, remained. Thus, the conclusions of the present section still stand: A principle involving the term ‘energy’, a principle which was regarded as definitional, or as analytic, if you please, has been abandoned. And its abandonment cannot be explained always as mere ‘redefinition’ or as change in the meaning of ‘kinetic energy’...

He continues:

But ‘All bachelors are unmarried’ cannot be rejected unless we change the meaning of the word ‘bachelor’ and not even then unless we change it so radically as to change the *extension* of the term ‘bachelor’. What makes the resemblance [to the ‘energy’ case] superficial is that if we were asked what the meaning of the term ‘bachelor’ is, we can *only* say that ‘bachelor’ means ‘unmarried man’, whereas if we are asked for the meaning of the term ‘energy’, we can do much more than give a definition. We can in fact show the way in which the use of the term ‘energy’ facilitates an enormous number of scientific explanations, and how it enters into an enormous bundle of laws. (Putnam 53)²⁰

But let’s consider the example of ‘bachelor’ more closely. It is still commonly assumed and asserted in philosophy that: it is analytically true by virtue of the meaning of ‘bachelor’ that:

X is a *bachelor* iff_{df} x is an unmarried adult male.

But this claim only seems plausible if one assumes that heterosexuality is universal, or that there is no way other than marriage for one to enter into a formalized lifelong commitment. It seems

plausible to say that an unmarried gay man who has made a lifelong commitment to another—perhaps even formalized it as a “civil union”—is not a bachelor. To press further: Is it analytic that marriage is between a man and a woman, or is it only “deeply embedded collateral information” (Putnam 1975a, 41)? Whose intuitions about ‘marriage’ should settle this?

One might suggest that a weaker claim is analytic by virtue of the meaning of ‘bachelor’:

If x is a *bachelor*, then x is unmarried.

But the truth of this claim depends on what sorts of institutions might qualify as “marriage,” and this is an issue that is highly contested and historically complex. For example, marriage as we know it has traditionally combined an economic institution with a quasi-religious institution setting constraints on sexual behavior. This is, of course, not an accident, since sex tends to produce offspring and offspring are, at least potentially, both an economic drain and an economic resource. However, it is possible to imagine a case in which the economic institution of marriage and the sexual institution of marriage are separated to form two kinds of marriage, a sexual marriage and an economic marriage. (Consider, perhaps, a variation on Margaret Atwood’s, *The Handmaid’s Tale*.) Further, suppose that one can be sexually married to A and economically married to B. I’m inclined to think that bachelorhood is really about sexual availability, so the fact that a man is economically married to A does not compromise his bachelor status, since he is still available to be sexually married to someone else.

In any case, the reliance on a background social framework is apparent in the case of:

X is a *parent* iff_{df} x is an immediate progenitor.

In some social/historical contexts this may seem analytically true by virtue of what ‘parent’ means. But laws and customs change so that one can become the legal parent of a child who is not biologically related, and with time, such parents are recognized as “real parents”.

In the case of *parents* and of *marriage* there are competing models of social life, of what’s essential and what’s accidental to our existing social structures. But just as what is essential or accidental to being an atom or being energy will depend on the background physical theory in which the term ‘atom’ is used, so what is essential or accidental to being a parent, or being

married, or being a bachelor, will depend on the background model of social life. (See also Burge 1986.) Putnam suggests that we should contrast cases such as 'bachelors are unmarried men' with scientific principles because the former is as close to pure stipulation as we can get, and the latter have "systematic import". Because of the systematic import of scientific principles, we can give up one or another of them without changing the meaning of the terms used to express them. (Putnam 1975a, 40) What Putnam (and others) seem to miss is that 'bachelors are unmarried men' also has systematic import, i.e., the concepts, in particular, of *marriage*, *adult*, and *male*, although familiar from common parlance, can also be the subject of social and political theory and of social contestation. Although the scientific essentialists were apt to claim that the analytic/synthetic distinction stood in the way of scientific progress, they were not as apt to see that it may also stand in the way of social progress.

What these cases reveal is that often what we take to be analytic principles actually encode certain social arrangements, and the relationship between terms encode certain power structures, e.g., the term 'parent' brings with it a certain normative weight, entitlement, etc. that the term 'primary caregiver' doesn't. Putnam suggests that it would be difficult to imagine a physical theory that did not employ some notion of *the past*, or of *energy*. These are framework concepts. Similarly it is difficult to imagine a social theory that did not employ some notions of *male*, *female*, *parent*, even something like *marriage*. So there is a reason why social constructionists want to rethink the term 'parent' rather than substituting the term 'primary caregiver'. And it is not surprising that in the United States we are fighting over *what the term 'marriage' means* as part of the struggle for gay rights. The term 'marriage' is a framework concept that links the institution to a broad range of other social phenomena, and does so in a way that 'civil union' cannot approximate.

The constructionist about 'parent' maintains that in cases where the manifest concept of *parent* is of an immediate progenitor, it may nonetheless be appropriate to understand the concept of *parent* as of *primary caregiver* (or some such notion). Is the constructionist simply changing the meaning of the term? If the concept of *parent* is a (social) framework concept,

which seems plausible, and if the work we need the term 'parent' to do is no longer best served by assuming that parents are immediate progenitors, then it is reasonable to consider this claim, not as giving the meaning of the term, but as stating a useful, though not universal, generalization.

This, then, suggests a reason to prefer the constructionist to the error theoretic strategy in analyzing at least some social kinds. The conceptualist's insistence that the concept in question, say *parent*, should be analyzed only in terms of what is manifest to us, can have the effect of fossilizing our social structure: if we are not allowed to adjust the contents of our framework concepts in light of developments in social theory and social life, then social change will require a wholesale adoption of a new conceptual scheme. Given that this is unlikely, change will be difficult. Moreover, because framework concepts are embedded with normative principles, rejecting the concepts may leave us with old practices and no new principles to guide us. If we combine the conceptualist strategy with the retention of purportedly natural categories, we further entrench the existing framework by suggesting that its analytic structure is just tracking nature's joints, not ours. So, in effect, the constructionist is making two moves that potentially destabilize our social arrangements: revealing that a purportedly analytic statement is in fact contingent generalization, and revealing that a natural category is in fact social.

8. Conclusion

I started this essay by asking whether social constructionist analyses of familiar terms or concepts can ever be philosophically acceptable if such analyses aim to *debunk* our ordinary understandings and so inevitably violate our intuitions. Further, if social constructionist analyses can be counterintuitive, are any limits on how counterintuitive they can be and still be acceptable; more generally, what makes for goodness in a constructionist analysis, if intuitions don't matter? And finally, are there any reasons why constructionist analysis should be preferred over error theoretic accounts that remain committed to the manifest naturalizing

concept as “what we mean” and offer a new social concept to capture how the term operates in our practice?

I have argued that constructionist analyses cannot be faulted *in general* for changing the subject, or for being counter-intuitive. Semantic externalism allows us to claim that what we are talking about is, in fact, a social category, even if we think it isn't. And in the case of framework concepts, social theory and social life may lead us to reject principles that seemed definitional, while we still retain the concept. So the constructionist can claim to provide an acceptable analysis of a concept, even if it is not intuitive.

A successful constructionist debunking will be one in which the best account of what we are doing (or should be doing) in drawing the distinctions in question—taking into account what legitimate purposes are being served and what objective types there are—has us tracking a social type.²¹ Such an account will not be purely a priori and will draw on social theory. I have not, in this essay, argued that a social constructionist account of race is preferable to an error theoretic or naturalistic account (though I have maintained that it is, see Haslanger 2000); such an argument would have to delve into empirical matters that I haven't touched on here.

It would be a mistake to conclude from what I've argued that constructionist analyses are always preferable to error theoretic analyses when there is a gap between manifest and operative concepts. Cases have to be examined individually. In particular, the arguments I've offered in this essay only address error theorists who rely on a conceptualist picture, i.e., those whose analysis of the controversial concept rests entirely on balancing intuitions. There may be some who adopt an error theory as the result of a broad analysis of our practices and purposes. My arguments do not weigh against such an account. Rather, I have urged that if our manifest concepts are misleading about our practices and mask what we are really doing with our concepts, we should consider whether there is a story to be told about how and why. If there is such a story, our accounts should reflect it.

Grounding philosophical analysis in linguistic competence or a priori intuition concerning our manifest concepts risks perpetuating social self-deception.

Although we cannot proceed without intuition, neither can we proceed without critical social theory. My hope is that the example of gender and race will encourage philosophers to pay greater attention to the rather ubiquitous gap between manifest and operative concepts, leading to less focus on our intuitions and more on the role of concepts in structuring our social lives. Philosophical analysis has a potential for unmasking ideology, not simply articulating it.

REFERENCES

- Alcoff, Linda. 2000. "Is Latino/a Identity a Racial Identity?" in *Hispanics and Latinos in the United States: Ethnicity, Race, and Rights*, ed., Jorge J. E. Gracia and Pablo De Grieff. New York: Routledge, pp. 23-44.
- Anderson, Elizabeth. 1995. "Knowledge, Human Interests, and Objectivity in Feminist Epistemology," *Philosophical Topics* 23(2):27-58.
- Andreason, Robin. 2000. "Race: Biological Reality or Social Construct?" *Philosophy of Science* 67 supplementary volume, pp. S653-66.
- Appiah, K. Anthony. 1996. "Race, Culture, Identity: Misunderstood Connections. " In A. Appiah and A. Gutmann, *Color Conscious: the Political Morality of Race*. Princeton: Princeton University Press, pp. 30-105.
- _____. 1993. *In My Father's House*. New York: Oxford University Press.
- Armstrong, David. 1989. *Universals: An Opinionated Introduction*. Boulder, CO: Westview Press.
- Burge, Tyler. 1986. "Intellectual Norms and Foundations of Mind." *Journal of Philosophy* LXXXIII(12): 697-720.
- _____. 1979. "Individualism and the Mental." *Midwest Studies in Philosophy* 4: 73-121.
- Glasgow, Joshua. 2003. "On the New Biology of Race" *The Journal of Philosophy* 100: 456-74.
- Gooding-Williams, Robert. 1998. "Race, Multiculturalism and Democracy." *Constellations* 5(1): 18-41.
- Hacking, Ian. 1999. *The Social Construction of What?* Cambridge, MA: Harvard University Press.

- Hardimon, Michael O. 2003. "The Ordinary Concept of Race." *Journal of Philosophy* 100:9 (September): 437-55.
- Haslanger, Sally. 2006. "Race and Natural Kinds." Conference presentation January 9. *Revisiting Race in a Genomic Age*, Stanford Humanities Center.
- _____. 2005. "What are we Talking About? The Semantics and Politics of Social Kinds." *Hypatia* 20(4): 10-26.
- _____. 2003. "Social Construction: The "Debunking" Project," in *Socializing Metaphysics: The Nature of Social Reality*, ed., Frederick F. Schmitt. Lanham, MD: Rowman and Littlefield, pp. 301-325.
- _____. 2000. "Gender and Race: (What) Are They? (What) Do We Want Them To Be?" *Noûs* 34(1): 31-55.
- _____. 1995. "Ontology and Social Construction," *Philosophical Topics* 23(2): 95-125.
- Higginbotham, 1998. "Conceptual Competence." *Philosophical Issues* 9: 149-162.
- Kitcher, Philip. 1999. "Race, Ethnicity, Biology, Culture" in L. Harris *Racism*. New York, Humanity Books, pp. 87-117.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lee, Sandra Soo-Jin, Johanna Mountain, and Barbara Koenig. 2001. "The Meanings of "Race" in the New Genomics: Implications for Health Disparities Research." *Yale Journal of Health Policy, Law and Ethics* 1: 33-75.
- MacKinnon, Catharine. 1987. *Feminism Unmodified*. Harvard University Press.
- Mallon, Ron. 2004. "Passing, Traveling and Reality: Social Construction and the Metaphysics of Race" *Noûs*, 38(4): 644-673.
- Mills, Charles. 1998. *Blackness Visible: Essays on Philosophy and Race*. Cornell University Press.
- _____. 1997. *The Racial Contract*. Cornell University Press.
- Mountain, Joanna L. and Neil Risch. 2004. "Assessing Genetic Contributions to Phenotypic Differences Among 'Racial' and 'Ethnic' Groups." *Nature Genetics* 36(11 Suppl): S48-53.

- Omi, Michael, and Howard Winant. 1994. "Racial Formation," In M. Omi and H. Winant, *Racial Formation in the United States*, Ch. 4. New York: Routledge, pp. 53-76.
- Peacocke, Christopher. 1998. "Implicit Conceptions, Understanding, and Rationality," *Philosophical Issues* 9: 43-88.
- Putnam, Hilary. 1975a. "The Analytic and the Synthetic." In H. Putnam, *Mind, Language and Reality*. Vol 2 of *Philosophical Papers*. Cambridge: Cambridge University Press, pp. 33-69.
- _____. 1975b. "The Meaning of "Meaning"" . In H. Putnam, *Mind, Language and Reality*. Vol 2 of *Philosophical Papers*. Cambridge: Cambridge University Press, pp. 215-271.
- _____. 1973. "Meaning and Reference." *The Journal of Philosophy* 70: 699-711.
- Risch, Neil, Esteban Burchard, Elad Ziv, and Hua Tang. 2002. "Categorization of Humans in Biomedical Research: Genes, Race and Disease." *Genome Biology* 3(7): comment 2007.1-2007.12.
- Rosenberg, Noah A., Jonathan K. Pritchard, James L. Weber, Howard M. Cann, Kenneth K. Kidd, Lev A. Zhivotovsky, and Marcus W. Feldman. 2002. "Genetic Structure of Human Populations." *Science* 298: 2381-2385.
- Zack, Naomi. 2002. *Philosophy of Science and Race*. New York: Routledge.
- _____. 1997. "Race and Philosophic Meaning." In *RACE/SEX: Their Sameness, Difference, and Interplay*. New York: Routledge, pp. 29-43.

Thanks to Louise Antony, Lawrence Blum, Alex Byrne, Jorge Garcia, Richard Holton, Erin Kelly, Ishani Maitra, Mary Kate McGowan, Lionel McPherson, Laura Schroeter, Marion Smiley, Sarah Song, Ronald Sundstrom, Ásta Sveinsdóttir, Gregory Velazco y Trianosky, and Steve Yablo for discussing with me the issues raised in this paper. Earlier versions of this paper were given at the Society for Analytic Feminism, the Australian National University, The California Roundtable on Philosophy and Race (2005), the University of Colorado, Boulder, and Tufts University. I am grateful to the participants in the discussions for helpful feedback.

² In (Haslanger 2003) I contrast this sort of interpretation with one that is more common in the context of “the science wars” which is discussed at length, e.g., in Hacking 1999.

³ I’ll continue, however, to speak of those whose project it is to argue that a category is socially founded as “social constructionists”, both because “social foundationalists” would be a serious misnomer, and also because it is reasonable to cluster those who make a variety of different social arguments together, even if their views are not always compatible. It may be down the road that this acceptance of the “social constructionist” label is more trouble than it is worth.

⁴ The interdisciplinary debate over race naturalism and the relevance of “race” for medicine is substantial and complex. I have listed only a few of the most controversial articles as examples.

⁵ I will use upper-case terms such as “White” and “Black” for races, lower-case terms such as “white” and “black” for the “color” markings associated with the races.

⁶ Quine distinguishes different forms of definition, the third being what he calls (drawing on Carnap) “explicative”. In giving explicative definitions, “an activity to which philosophers are given, and scientists also in their more philosophical moments...the purpose is not merely to paraphrase the definiendum into an outright synonym, but actually to improve upon the definiendum by refining or supplementing its meaning.” (Quine 1963: 24-5) “Ameliorative” captures better than “explicative” the sort of project Quine is characterizing as especially philosophical; it is this sort of project that I’ve also called “analytical” (Haslanger 2000). Because “analytical” is commonly used to characterize Anglo-American philosophy in general, and because I’m attempting here to introduce a more fine-grained framework, using “ameliorative” rather than “analytical” will sometimes avoid ambiguity. It should be understood, however, that on my view, whether or not an analysis is an improvement on existing meanings will depend on the purposes of the inquiry.

⁷ In previous work I’ve dubbed this the “conceptualist” approach. However, I’ve been convinced by others, and by confusions in discussion, that it may be better described as an “internalist” approach in order to highlight the contrast with the underlying externalism assumed by the descriptive approach (described next in the text). The change is useful, however, there are different degrees and kinds of internalism and externalism and I am only using the terms suggestively and not precisely here.

⁸ On objective types see Armstrong 1989.

⁹ It might be useful to see this by analogy with other terminological developments in science. Although our understanding of and even our definition of ‘atom’ has changed over time, it is plausible that there is something worthwhile we have been and continue to be talking about.

¹⁰ I don’t mean to suggest here that there is only one manifest concept and only one operative concept. The manifest and operative concepts may vary from context to context.

¹¹ However, we often make mistakes in applying our manifest concepts. When we make a simple mistake, must we postulate an operative concept distinct from the manifest one? I’m not sure much hinges on this, but it seems to me that if we have a sparse theory of objective types a better approach would be to understand the practice as tracking a nearby (or the nearest?) objective type; if the nearby type is the type also determined by the manifest concept, we have coincidence. The operative concept will be, then, the concept that determines that type in terms that make the most sense in analyzing the practice.

¹² Given what I say in (Haslanger 2005), we could also consider it a genealogist strategy.

¹³ Of course, in the case of *parent*, it is not plausible to adopt a thoroughgoing error theory, since there are some people who do satisfy the conditions of being an immediate progenitor. But the internalist strategy would have us adopt a qualified error theory: all of our uses of the term 'parent' aside from the core biological cases are strictly in error because they pick out people who aren't really parents. In response to the charge that this would be inadequate in, for example, a school context, the conceptualist could also maintain that such an approach to the semantics does not entail that we retain the concept or terminology of *parent* in our school practices: perhaps we would do well to address our school memos using 'primary caregivers' or another term.

¹⁴ Although Ian Hacking is a constructionist, he concludes that constructionists cannot simply be making the debunking point that the category in question is socially unified rather than naturally unified because the claim would be "redundant". I think the distinction between the manifest and operative concept helps show why it need not be redundant and may be important. See (Hacking 1999, 39).

¹⁵ Interestingly, Appiah (1996) does consider a kind of descriptivist approach, and still concludes that there are no races. He does so, however, because he takes it as a constraint on the type that can be projected from the paradigms that it be a natural type, since this is part of the concept, given his explorations of the history of the concept. So he does not consider the possibility that the objective type we designate with the term 'race' is a social type. This seems to be a "mixed" approach that places, to my mind, too much weight on the history of the concept and does not adequately recognize objectivity of social kinds. The argument I provide in what follows against error theories will only address those who defend their view using a conceptualist strategy. Those pursuing an ameliorative strategy who come, ultimately, to the conclusion that an error theory is the best option, I do not address here.

¹⁶ I myself am not opposed to proposing new meanings and, more generally, undertaking revisionary metaphysics. Given the history of our language and our conceptual framework, it would be a miracle if we had landed upon the best framework to describe the world. I only care about what we do mean as a step in an inquiry into what we should mean. Nonetheless, I think the conceptualist is misguided, and it is worth pointing out why, since many are not so happy with radical revisions as I am.

¹⁷ Although I'm not endorsing the methods of ordinary language philosophy, the complexity of our use of words in different contexts is something ordinary language philosophers were well attuned to, and some of their methods and ideas are tremendously valuable for genealogy.

¹⁸ Because the terminology of "natural kind" is used in several different ways, it will be helpful to make a few distinctions. The term "kind" is sometimes used to classify substances, in the ordinary case, (physical) objects. Substances can be classified according to their essence; kinds consist of groups of objects with a common essence. E.g., tigers constitute a kind of thing because each tiger has essentially a certain cluster of properties that define the kind. On other occasions, the term "kind" is used to refer to what are sometimes called *types*. A type is a group of things, sometimes substances, but possibly non-substances, that has a certain unity. This unity need not be a matter of sharing essential properties: red things constitute a type (their unity consists in their all being red), even though redness is seldom an essential property of the things that have it. Unity seems to come in different degrees. The things on my desk might be thought to constitute a weak sort of type (they have in common the fact that they are on my desk), and at the limit there are highly gerrymandered sets of things that don't have any unity at all and so don't constitute a type.

One way to think about the unity of types is in terms of similarity between the members. We can distinguish different sorts of types by distinguishing axes of similarity. Exactly six foot tall human beings are a natural type because the commonality between the members is natural (species and height); high school graduates are a social type because the commonality between the members is social. Both of these types are (metaphysically) objective, however. How to draw the line between social and natural types is difficult (as is the distinction between objective and subjective!) and not one I will address here. I'll have to rely on background understandings and familiar cases. However, it is important to keep in mind that as I am using the terms the distinction between objective and non-objective kinds/types is importantly different from the distinction between natural and social kinds/types.

¹⁹ The third sort is supported by the idea that inquiry can seem to converge on an idea or concept that we seemed to have in mind all along, even though no one, even the best minds, could have explicated it. So, e.g., in the history of logic, math, and science, it is plausible to maintain that certain experts were

“grasping a definite sense, whilst also failing to grasp it ‘sharply’.” (Peacocke 1998, 50) I’ll assume that these can fall within objective type externalism because at least there are some paradigms that fix the reference in question; they don’t seem to qualify for social externalism because there isn’t yet a standard linguistic usage.

²⁰ Putnam’s paper is famous for going on to argue that there are scenarios in which we would consider evidence that ‘all bachelors are unmarried is false’; but such scenarios are ones in which ‘bachelor’ comes to function as a natural kind term for those with a certain neurosis. So again, on Putnam’s account the phenomenon of conceptual evolution occurs in the context of developing natural science.

²¹ On the issue of what counts as a “legitimate purpose” and how our theoretic purposes should be evaluated, I follow (Anderson 1995) and trends in feminist empiricism more broadly.