

On How to Achieve Reference to Covert Social Constructions

E. Díaz-León

Department of Philosophy, University of Barcelona

What does it mean to say that some features, such as gender, race and sexual orientation, are socially constructed? Many scholars claim that social constructionism about a kind is a version of realism about that kind, according to which the corresponding kind is a social construction, that it, it is constituted by social factors and practices. Social constructionism, then, is a version of realism about a kind that asserts that the kind is real, and puts forward a particular view about the nature of the kind, namely, that it is constituted by social factors and practices. Social constructivists about human kinds such as gender, race and sexual orientation often make an additional claim, namely, that these kinds are social constructions but they are typically believed to be biological kinds (that is, people are typically wrong about the nature of these kinds). Ron Mallon (2017) calls social constructions that are (falsely) taken to be biological kinds *covert social constructions*. This paper is about how we could have terms in our natural language that come to refer to covert social constructions.

Keywords: social construction, theories of reference, externalism, race, gender

What does it mean to say that some features, such as gender, race and sexual orientation, are socially constructed? Many scholars claim that social constructionism about a kind is a version of realism about that kind, according to which the corresponding property is a social construction, that it, it is constituted by social factors and practices. Social constructionism, then, is a version of realism about a kind that asserts that the kind is real, and puts forward a particular view about the nature of the property, namely, that it is constituted by social factors and practices.

Social constructivists about human kinds such as gender, race and sexual orientation often make an additional claim, namely, that these kinds are social constructions but they are typically believed to be biological kinds (that

Corresponding author's address: E. Díaz-León, Department of Philosophy, University of Barcelona, C/ Montalegre, 6, Barcelona 08001, Spain. Email: ediazleon@ub.edu.

is, people are typically wrong about the nature of these kinds).¹ Ron Mallon (2017) calls social constructions that are (falsely) taken to be biological kinds *covert social constructions*. This paper is about how we could have terms in our natural language that come to refer to covert social constructions.

Mallon (2017) presents a challenge to the view that our terms for race, gender, sexual orientation and other allegedly socially constructed kinds can refer to covert social constructions. In what follows I will introduce the challenge, and I will develop and defend a solution to the challenge on behalf of the social constructionist that can meet Mallon's objections, or so I will argue.

Mallon's challenge to the claim that terms such as "race," "gender" and "sexual orientation" refer to covert social constructions is the following, in a nutshell: no matter what theory of reference we assume, there will be problems for the claim that those terms refer to covert social constructions. In order to illustrate this point, I will explain the problem for the descriptivist theory of reference and for the causal theory of reference.

First, the descriptivist theory asserts that a term *T* refers to an entity *X* if and only if *X* satisfies the descriptions associated with term *T*. For this reason, if we assume the descriptivist theory of reference, then terms such as "gender," "race" and "sexual orientations" will refer to covert social constructions only if these satisfy the descriptions associated with those terms. But, and this is the crucial idea, if these are *covert* social constructions, this is because ordinary speakers are typically wrong about their nature, that is, they typically believe that the referents of those terms are not social constructions, and in particular they take them to be biologically significant kinds. Therefore, it is difficult to see how those terms could refer to covert social constructions, since this assumes that ordinary speakers do not associate descriptions with those terms that capture the nature of the referent, since the descriptions they do associate typically say that the referents are biologically significant kinds, but it is clear that covert social constructions at issue here are not biologically significant kinds, and therefore they do not satisfy the corresponding descriptions. Hence, a descriptivist theory of reference cannot make sense of the claim that those terms refer to covert social constructions (see Appiah 1996, Glasgow 2009).

In response to this worry, some social constructionists have appealed to the causal-historical theory of reference in order to argue for the possibility that the corresponding terms refer to covert social constructions (e.g., Haslanger 2006). According to the causal-historical theory of reference, a term *T* refers to an entity *X* if and only if *X* is at the beginning of a causal chain of successive uses of the term, where the causal chain was originated

¹ See for instance (Haslanger 2006), (Glasgow 2009) and (Stein 1999).

with a baptism or dubbing of X as T . Even if some information about X may be necessary in order to identify X so that the original speakers can associate it with the term T , this term can then be passed on from speaker to speaker without the need for these speakers to associate any identifying information with the term, they just need to be part of the corresponding causal chain (and at most, to have the intention to refer to the same thing that previous speakers are referring to). Because of this, social constructionists argue, terms such as “gender” and “race” might refer to properties such as covert social constructions even if ordinary speakers are wrong about the nature of such referents, and in particular even if they believe falsely that they are biologically significant kinds, since speakers do not need to associate identifying descriptions with the term for the term to refer successfully. So the causal-historical theory allows the possibility that the corresponding terms might refer to covert social constructions after all, in spite of the fact that many ordinary speakers believe that gender and race are biologically significant kinds.

However, this view also faces difficulties, according to Mallon. The crucial idea is this: according to social constructionism, covert social constructions are properties that are partially constituted by social practices, and in particular, a social construction K is partially constituted by the social practices of labelling and classifying some individuals as K and treating them in certain ways because of being perceived as K . For this reason, Mallon argues, if we go back to the initial baptism when a certain term T was introduced, at that time the covert social construction could not have existed, precisely because the kind is partially constituted by the term T itself, so there could not be a preceding kind K for the term T to pick out. This is why appealing to the causal-historical theory of reference is problematic for social constructionists. Mallon calls this the *missing referent* problem: if covert social constructions are supposed to be partially constituted by the employment of term T itself, then there could not be a social construction of that kind prior to the introduction of term T , and then the original baptism could not have occurred, so appealing to the causal-historical theory is moot.

Mallon discusses two possible responses to this problem, namely, the idea that there was a referent after all prior to the original baptism, and the idea that even if there was no referent of the right kind at the time of the original baptism, the referent of the corresponding terms has switched, so that the terms have now come to refer to covert social constructions, as social constructionism has it. Mallon argues that none of these responses can be successful. In response, I will argue that they can both work, and I will provide a version of each option in order to show that there can be successful defenses of social constructionism.

1. Finding the missing referents

Can we say that there was a referent prior to introducing term *T*? And could that referent be a covert social construction? Mallon mentions two possible referents that could exist prior to introducing term *T*, namely, what he calls *thin natural kinds*, and *institutional kinds*. Let me elaborate. Thin natural kinds are superficial kinds, that is, kinds that appear in nature without the need of human intervention (although there might also be thin natural kinds that are created by humans), and such that they are not explanatorily robust, that is, they do not appear in many law-like generalizations. For instance, kinds such as having such and such skin color, or such and such hair texture, count as thin natural kinds. We can assume that these kinds were instantiated prior to introducing the term “race,” so the term could have picked out one of those thin natural kinds. But it is clear, Mallon argues, that this is the *wrong kind of referent* for the social constructionist, since thin natural kinds are different from covert social constructions: whereas social constructions are constituted by social practices, thin natural kinds are not. At most, thin natural kinds could have been *caused* by social practices, but they are not *constituted* by social practices, so they cannot be covert social constructions of the sort social constructionists identify with race or gender.

What about institutional kinds? Institutional kinds are, roughly, social kinds that have been created by explicit conventions, such as *money*, *Prime Minister*, *Supreme Court*, and *university*.² Again, Mallon argues that institutional kinds are the wrong kind of referent, since covert social constructions are such that they are typically taken (falsely) to be biologically significant kinds, whereas institutional kinds are not usually taken to be biologically significant kinds. As Mallon argues, ordinary speakers are typically aware of the nature of institutional kinds, since institutional kinds are created by explicit assignments of functions, whereas covert social constructions are supposed to be kinds such that ordinary speakers are wrong about their nature.

I want to argue that there are further options for the social constructionist. In particular, a social constructionist that appeals to a causal-historical theory of reference can explain how social kind terms can refer to covert social constructions because there could be social constructions prior to introducing the corresponding terms that are neither thin natural kinds nor institutional kinds. In particular, the social constructions at issue here could be a matter of social constructions that are not constituted by the employment of term (or concept) *T*, but they are nonetheless constituted by social

² See (Searle 2009) and (Guala 2016), among others, for further discussion about the nature of institutional kinds.

practices that do not involve employment of the term. Here it will be useful to distinguish between social constructions that are constituted by social practices in general, and social constructions that are constituted (at least in part) by the employment of concepts or terms that aim to refer to those very social constructions.³ Examples of the latter are kinds such as “Prime Minister” and “Supreme Court”: these social kinds are partially constituted by the employment of the corresponding terms. It can be argued that when the terms are introduced, the corresponding social kinds are thereby created. The introduction of the term and the bringing about of the social kind occur at the same time, and this is how the social kind can be partially constituted by the employment of the very same term. But as we have seen, these kinds are very different from covert social constructions. Indeed, these social kinds that are partially constituted by the very terms are typically *overt* social constructions, so they cannot be of help here.

But not all social kinds are constituted by the very same concepts and terms that are introduced when the social kinds are created. In particular, there are social kinds that are constituted by social practices that do not involve terms or concepts that are introduced with the intention of referring to them and because of this, social kinds of this sort could have existed prior to the original baptism. For instance, social kinds such as *inflation* or *racism* could exist before the corresponding terms “inflation” and “racism” were introduced (Thomasson 2003). That is, these social kinds are constituted by complex social practices that do not involve the employment of those very terms or concepts (but they probably involve the employment of many other concepts). For this reason, we can imagine ceremonies of original baptism where these terms were introduced, so that they came to refer to inflation and racism given that these social kinds existed prior to the naming ceremonies.

Can we say something similar about gender and race? The view would be as follows: gender and race would have to be identified with social kinds that do not involve the employment of the terms “gender” and “race” themselves. What would these social kinds look like? I will focus on the case of “race.” We could argue that before the term “race” was introduced, there were already social constructions in the vicinity, namely, there were groups of people with certain ancestors in common, or with certain geographical origins, that were treated in a privileged or a subordinated manner in virtue of having those features. And this is in itself a socially constructed kind. That is to say, the social kind at issue here is not the property of being treated in a certain way because of being classified as a member of a certain *race*. This social kind is indeed partially constituted by the concept of race itself.

³ See (Díaz-León 2018) for further discussion of this distinction.

But my view appeals to a slightly different socially constructed property, namely, the property of being treated in a privileged or subordinated manner because of being perceived to be part of a certain geographical group, or because of (being perceived as) having certain skin color, or having certain ancestors in common, or some other superficial features.⁴ None of these social practices presupposes the employment of *racial* concepts. Given the prior existence of these socially constructed kinds, we can now understand what is going on at the original baptism ceremony. The introducers of the term “race” might have had the intention of referring to those groups that were treated in such and such manner because of their geographical origin (or some other superficial feature), but they might have the additional (false) belief that such treatment was justified because the members of each group also shared certain underlying biological properties that justified the privilege or discrimination.⁵ This explains why ordinary speakers associate race terms with false beliefs, namely, the belief that races are biologically significant kinds. But the causal-historical theory of reference does not require that ordinary speakers have correct beliefs about the referent: it is enough that they are part of the right causal chain, and that there was a successful act of original baptism.

Someone could wonder how the original baptism could be successful, if the original speakers had false beliefs about the alleged referents (to wit, the referents were actually covert social constructions whereas the speakers who introduced the terms falsely believed that they were biologically significant kinds). The key condition here is that there is a kind in the vicinity of the original baptism that was explanatorily robust enough so that it could become the referent of the term. And it seems to me that when the original speakers introduced the term with the intention of referring to certain groups of people with such and such skin color and such and such geographical origins (or whatever was the relevant thin natural kind), whom they also took to share biobehavioral essences that justified the differential treatment at the time (but did not really instantiate these underlying essential properties), then the referents could very well be the existing socially constructed

⁴ See (Díaz-León 2015) for a discussion of a similar social constructionist account of race.

⁵ In my view, this false belief can explain why the introducers of the term “race” were motivated to introduce a new term, in addition to the terms they already have for the thin natural properties: there were groups of people with thin natural properties in common who were privileged or discriminated in virtue of having those thin natural properties, and in addition to this some people believed that such privilege or discrimination was *justified* because those groups shared an underlying property in common, such as a biological essence. This justificatory aim is what explains the introduction of the new term “race,” which happened to refer to those previously existing social properties, about which people had false beliefs.

properties of the sort I have described, since these are the most explanatorily useful kind in the vicinity.⁶

A remaining worry is the following. It might be argued that on this view, when racial terms are introduced for the first time, the existing social properties in the vicinity already seem to be constituted in part by *racial* concepts, since people had ways of describing those groups such as “groups of people who shared thin natural properties such as phenotypic features or geographical origins or common ancestors, that are privileged or discriminated in virtue of having those properties,” which already seems to encapsulate a *racial* concept. This description seems very similar to the concept of race that is introduced in the baptism ceremony according to this view, so why are those social properties not constituted by racial concepts. My answer to this worry is twofold. First, it could be argued that at the moment of the baptism ceremony, the introducers of the term had the additional (false) belief that the individuals of the groups also shared an underlying essential property, such as a biobehavioral essence, that justified the differential treatment. This new belief might explain why there arises a new concept, although this response is a bit tricky since strictly speaking the causal theory of reference has it that the beliefs of the speakers are not part of the content of the concept (nor the meaning of the term). Therefore, the presence of this essentializing (false) belief would not make a difference to the content of the concept, and then we cannot explain why the baptism ceremony brings about the introduction of *racial* concepts for the first time. (That is to say, this response might be compatible with a hybrid version of descriptivism that incorporates elements of the causal theory, but not with the causal theory of reference itself.)

A second response that is more akin to the causal theory of reference might be the following. Mark Sainsbury and Michael Tye (2012) have argued that concepts are individuated by their historical origins, not in virtue of their semantic or epistemic content. So, two concepts might be distinct if they have different origins, even if their semantic or epistemic content is the

⁶ Guido Löhr (2019) has similarly argued, in response to Mallon (2017), that our social kind terms can refer to covert social constructions only if the covert social kinds are not constituted by concepts aiming to refer to those very same social kinds. Löhr also claims that this makes these social kinds mind-independent enough so that semantic externalism can explain reference to them, but, as he acknowledges, these kinds are social constructions precisely because they are constituted by social practices, including the employment of many concepts and the entertainment of many beliefs, other than the concept referring to the social kind in question, as I explain above, so they cannot be mind-independent strictly speaking. Löhr (2019) also develops a very interesting account of how covert social kinds can be stable enough so as to be the referents of our terms via the causal theory of reference.

same. According to this view, then, any concept referring to the relevant social groups that existed before the introduction of racial terms, and the concept referring to the very same social groups expressed by the term “race” introduced in the baptism ceremony, would be different concepts, to the extent that they have different historical origins, even if they are co-referential.

Hence, we have two different ways of distinguishing the relevant concepts. First, on a view more akin to the descriptivist theory of reference, the concepts are different because the associated beliefs are different. Second, on a view more akin to originalism, the concepts are different because they have different historical origins.

2. Reference switch

Alternatively, could we say that the referent of terms such as “race” has *switched* over time? This is also a coherent possibility, or so I want to argue. There might be cases such that there were no socially constructed properties that were explanatorily robust enough so as to become the referents via the original baptism. In these cases, there might be thin natural kinds in the vicinity, and the corresponding gender or racial terms might come to refer to those, via the corresponding original baptism ceremonies. And then, after time, the individuals that belonged to those thin natural kinds might have acquired new socially constructed properties, namely, socially constructed properties that involve the employment of the corresponding terms or concepts themselves. For instance, someone could argue that the term “race” originally referred to thin natural kinds, but that after a while, given the social practices of classifying people into racial groups and treating them with privilege or discrimination in virtue of their racial group, then members of these thin natural kinds acquired new properties, namely, the socially constructed property of being treated in a different way because of being taken to belong to a racial group (that is, to fall under a racial concept). And it could be argued that once these socially constructed properties became explanatorily robust enough, they could have come to be the referents, via reference switching, precisely because they became more explanatorily robust than the thin natural kinds (since these thin natural kinds are not very explanatorily robust, as we have explained above).

But how could this mechanism of reference switching work? According to Mallon, a clear case of reference switching was described by Gareth Evans (1973), where he explained the case of “Madagascar.” This name used to refer to a region of the mainland of Africa, but apparently, when Marco Polo heard the natives use the name, he thought they were referring to the island that is now called “Madagascar,” and the new use spread around. This is a case of reference switching where Marco Polo had the intention of referring to

the same entity as the previous speakers, but he had different beliefs that did not match with the previous referent (such as “being an island”), and these new beliefs somehow trumped the intention to refer to the same entity as the previous speakers, and brought about a new grounding for the name, so that it came to refer to a new entity. According to Mallon, this is a crucial feature of reference switching, namely, there have to be *new beliefs* that change the referent. But in the case of the view I sketched above, where racial terms used to pick out thin natural kinds but they now pick out covert social constructions, there has been no change of beliefs, precisely because these social constructions are *covert*, that is, ordinary speakers typically believe (wrongly) that races are biologically significant kinds, and keep associating this belief with the term “race.”

Therefore, this view claims that the term “race” used to pick out thin natural kinds, and it now picks out covert social constructions, where neither of those is a biologically significant kind, although ordinary speakers still associate the (false) belief that the referents are biologically significant kinds. What brought about this change of referent then, if there was no change of beliefs whatsoever? It is true that this case is different from the Madagascar case as discussed by Evans. In the racial case, according to my suggestion, there is no change of beliefs: ordinary speakers keep associating false beliefs with the term “race,” namely, the belief that individuals who belong to the same race share biobehavioral essences. Hence, what brings about the change of reference is precisely the apparition of a more explanatorily robust social kind in the vicinity, namely, the socially constructed kind that is partially constituted by the employment of the term “race” itself. Given reference magnetism, the terms that fix their referents by means of a causal theory of reference have to refer to the most natural property in the vicinity that captures the patterns of use of the term. If a new, more explanatorily robust kind is created that equally well explains the patterns of use of the term, this can become the referent of the term.⁷

As we have seen, then, there are good reasons to think that our gender and racial terms might refer to covert social constructions after all. Either they came to refer to covert social constructions that existed prior to the original baptism ceremonies, or the terms came to refer to covert social constructions via reference switching. We can therefore conclude that Mallon does not pose good reasons to think that any of these options is untenable. So covert social constructionism can live another day.

⁷ See (Sider 2017) for an interesting discussion of reference magnetism with respect to the reference-fixing of gender and racial terms and other social kind terms.

Bibliography

- Appiah, K. A. (1996). Race, culture, identity: Misunderstood connections, in K. A. Appiah and A. Gutmann (eds), *Color Conscious: The Political Morality of Race*, Princeton University Press, pp. 30–105.
- Díaz-León, E. (2015). In defence of historical constructivism about races, *Ergo* **2**: 547–562.
- Díaz-León, E. (2018). Kinds of social construction, in P. Garavaso (ed.), *Bloomsbury Companion to Analytic Feminism*, Bloomsbury Academic, pp. 103–122.
- Evans, G. (1973). The causal theory of names, *Proceedings of the Aristotelian Society* **47**: 187–208.
- Glasgow, J. (2009). *A Theory of Race*, Routledge.
- Guala, F. (2016). *Understanding Institutions: The Science and Philosophy of Living Together*, Princeton University Press.
- Haslanger, S. (2006). What good are our intuitions? Philosophical analysis and social kinds, *Proceedings of the Aristotelian Society* **80**: 89–118.
- Löhr, G. (2019). Social constructionism, concept acquisition and the mismatch problem, *Synthese* pp. 1–15.
- Mallon, R. (2017). Social construction and achieving reference, *Nôus* **51**: 113–131.
- Sainsbury, M. and Tye, M. (2012). *Seven Puzzles of Thought and How to Solve Them: An Originalist Theory of Concepts*, Oxford University Press.
- Searle, J. (2009). *Making the Social World*, Oxford University Press.
- Sider, T. (2017). Substantivity in feminist metaphysics, *Philosophical Studies* **174**: 2467–2478.
- Stein, E. (1999). *The Mismeasure of Desire: The Science, Theory and Ethics of Sexual Orientations*, Oxford University Press.
- Thomasson, A. (2003). Realism and human kinds, *Philosophy and Phenomenological Research* **67**: 580–609.