Criticizing Social Reality from Within: Haslanger on Race, Gender, and Ideology

Stahl, Titus

TITUS STAHL

CRITICIZING SOCIAL REALITY FROM WITHIN
HASLANGER ON RACE, GENDER, AND IDEOLOGY

1. Introduction

Any kind of socially progressive critique of social practices must accomplish the difficult task of taking up a stance that is both appropriately critical of, and sympathetic to, the self-understanding of those whom it addresses. In doing so, it must avoid two mistakes: on the one hand, it must take into account that many of the negative features of our societies, such as racism and sexism, are not only rooted in what people do, but also pervade the very *conceptual categories* in which we understand ourselves. Thus, any serious critique of our social world has to seemingly reject many aspects of this socially dominant self-understanding. It seems necessary to do so in order to avoid falling into the trap of unconsciously reproducing relationships of oppression or subordination by formulating one's criticism in a language that already buys into a problematic conceptual framework. On the other hand, it is a futile enterprise to try to *completely* reject and replace the categories of the self-understanding of those whom one addresses. This is not only because it is hardly possible for critics to step completely outside of the language and the ways of thinking which are prevalent in their society. If they attempt to do so, critical theorists might also become unable, firstly, to correctly identify the subjective

experiences of oppression to which critical theories must necessarily refer; secondly, they risk becoming unable to formulate normative principles to which those whom they address could reasonably agree; and, thirdly, they become less capable of understanding the social struggles of their times. In other words, if social critics do not aim at an understanding of social practices from the 'inside' (that is, as understood from within those frameworks of thought and action that they aim to criticize), their critique becomes too disconnected to be valuable.

Even though one might entertain the thought, for example, that it would be better if we all just rejected the idea of 'race' and treat the concept of 'race' as having no meaning at all, it quickly becomes clear that not only would this ignore the fact that categories of 'race' are real for all of us, but that for many members of our societies the fact that they belong to one race and not to another is encountered each day as a brute fact. The experience of belonging to a 'race' cannot be understood without reference to the reality which 'race' actually has in our societies, understood in terms of the meaning of the relevant discursive ascriptions.

Critical theorists have traditionally employed a methodological solution to this dilemma, namely, the method of 'immanent critique'. As a method, immanent critique begins from the self-understanding of a given society and critically evaluates this self-understanding on its own terms in order to emphasize the ways in which it fails to successfully structure the practice of that society and to point out the pathologies that it necessarily produces. If one of the tasks of critical theories is to make this burden of a self-understanding being deficient according to its very own standards 'still more oppressive by adding to it a consciousness of it, and the shame […] more shameful by making it public', as Marx (1972, 134) famously argued, they might be capable of breaking the spell of ideological self-understandings without resorting to an 'external standpoint'.

While many social theorists engaged in progressive politics would agree to this description, the enormously difficult task of articulating this idea in the terms of our most advanced philosophical theories of language, mind, and social reality has rarely been attempted. It is the great achievement of Sally Haslanger's essays in *Resisting Reality* that she offers exciting new

answers on how to think about these problems. This especially concerns her analysis of the role of race and gender concepts, an analysis that neither rejects them in favour of some idealized notion of how we *better should* think or speak, nor accepts them as a mere given to which we have to adapt.

There are two points in *Resisting Reality* in which both the advantages and the challenges of this strategy become explicit. The first point concerns the treatment of the categories of race and gender (categories for which the problem that I have just described is especially salient). The second point concerns the notion of 'ideology'. I would like to argue that in both cases Haslanger's treatment of the issue goes beyond traditional approaches and establishes extremely interesting results. However, I would also like to examine both cases to see whether her approach involves a certain residual individualism that makes some of her arguments less powerful than they could otherwise be.

2. Semantic Externalism and Social Kinds

One of the most fundamental claims of *Resisting Reality* concerns a combination of *realism* and *constructionism* in regard to gender and race. Haslanger argues that concepts such as 'woman' or 'black' describe something *real* (as opposed to a mere illusion), but what they describe are not (as some people often think) essential or even natural properties of individuals, but rather *social kinds*. In particular, such terms describe the membership of people in groups that occupy certain positions in social hierarchies of domination (Haslanger 2012, 229-235).[1] With this claim, Haslanger not only rejects *naturalist* theories of race and gender, but also *eliminativist* theories that hold that race and gender terms do not refer to anything at all (cf. 299 ff.).

This theory rests on a certain picture about the meaning of concepts. Of course, Haslanger acknowledges that the intuitions that typical language users have about the meanings of race and gender terms do not necessarily go along with an explicit understanding that they refer to social kinds.

In fact, many people who talk about race and gender believe the corresponding terms to refer to *natural kinds*. However, drawing on externalist theories of meaning, Haslanger argues that an introspective analysis of concepts (that is, an analysis that refers to the understanding of ordinary language users) is inadequate to determine their content (379, 398). Instead, she endorses a variety of *semantic externalism*, that is, a view about concepts that holds that what determines their extension – and also, at least in part, their meaning – is what these concepts really 'track'.

We can therefore distinguish between two aspects of a concept: the *manifest* and the *operative concept* (92, 370). The manifest concept is determined by the meaning that language users *understand* a term to have. In contrast, the operative concept is determined by the properties or entities that are *actually tracked* by the linguistic practice in which such terms are employed.

As is obvious in the case of race and gender categories, there can be a mismatch between manifest and operative concepts. For instance, many people believe that these terms track essential, intrinsic properties of individuals that serve to explain their behaviour, whereas closer inspection reveals that there are, in fact, no such intrinsic properties. Rather, the way we employ categories of race and gender in our 'everyday theories' suggests that these concepts reliably track *positions of social status* due to which certain groups of people, as categorized by certain physical 'markers', are systematically treated differently.

This mismatch between operative and manifest concepts in our everyday use has direct implications for the question of critique: given this distinction, we can attempt to provide an immanent critique of our linguistic practices. This is because we are in principle capable of finding out (and convincing others) that our concepts of race and gender are not adequately understood using an essentialist theory of these phenomena. Such an immanent approach that points out a problem of a social practice (in this case, of the practice of categorizing persons according to gender) 'from within' is preferable to a critique 'from the outside' which only takes up the manifest concept and subjects it to a 'detached' metaphysical scrutiny. In this case, such an 'external' critique could only discover that

nothing in the real world is captured by the manifest, essentialist concept of race (because there are no racial essences that play any role in the best explanation of the behaviour of individuals). Consequently, an external critique of our 'race' vocabulary could only argue for *discarding* this vocabulary and replacing it with another. But such a critique not only forecloses the possibility of investigating the possible rationale of some of our discursive practices in which we employ this vocabulary, it also does not tell us anything about the best understanding of the experiences of those persons who are members of the respective social kinds.

This way of spelling out an 'immanent' strategy relies on a distinction between the semantic self-understanding of people and the actual social kinds that their concepts track that is often presented as relatively clean-cut. Even if Haslanger acknowledges that there can be multiple, competing concepts on both the operative and the manifest level (see 370, n. 5), she seems to clearly distinguish two possible types of social critique: first, a *critique of concepts* that shows that some concepts which many people take to refer to natural kinds actually refer to social kinds, and second, a *critique of society* that aims to change what social kinds there are. While the first type of critique might be a precondition for the second (we must know, for example, that 'woman' refers to a social kind in order to be able to intend to change society so that women, understood in terms of social status positions, no longer exist), these activities seem relatively independent from one another.

To see why this might be a problematic way to put it, it is useful to examine in more detail how different kinds of semantic 'externalism' conceptualize the 'external'. Haslanger introduces semantic externalism in two of its forms: first, there is the 'natural kind' variety which describes the view that the meaning of natural kind terms is partly constituted by the 'external' instances of that kind that are responsible for our use of that concept.[2] Second, there is a 'social variety', that is, the view that the meaning of some of our concepts is constituted by the linguistic usage of our community.[3] After referring to these two classical models, Haslanger introduces her own account that she calls 'objective type externalism'.

According to objective type externalism:

'Terms or 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.' (374)

It is clear that this is an extension of natural kind externalism to social kinds. In other words, it takes the *objective type* which is tracked by the concept to be central for its meaning. Of course, in contrast to traditional natural kind externalism, this account also allows for the possibility that the objective type in question is a social kind. Compared with social externalist approaches (which also apply to more than natural kind terms), however, the *communal use* of a term does not play any discernible role for its meaning on this account.

My main worry regarding this move towards a *specific* kind of externalism is not one about its independent plausibility as a position within the philosophy of language. In other words, I am not so much concerned with the success of 'objective kind externalism' as a theory of meaning, but rather with the consequences of choosing one such theory for a *critique* of concepts.

As a starting point, we could interpret Haslanger's treatment of the matter as entailing that our individual understanding of linguistic terms, as captured by the 'manifest concept', is fully independent from our social conventions which are constitutive for the social kinds that these terms track, that is, as an 'individualist' variant of objective social kind externalism. Such an interpretation could emphasize some useful features of objective kind externalism for her critical project that aims to uncover how individuals' understanding of the meaning of their concepts can mislead them. That is, the resulting account of meaning would support a type of philosophical critique that uncovers the ways in which beliefs of individuals about the proper use of race and gender terms only make sense by relying on assumptions which do not withstand philosophical reflection. But sharply dividing the semantic self-understanding of individuals (as the

basis of the manifest concept) from social reality (as constitutive for the social kinds which determine the operative concept) also runs the risk of concealing more 'material' aspects of the discourses in which these concepts are used.

This point might best be illustrated by looking at another theory of conceptual content: Robert Brandom's inferentialism. Even though Brandom's peculiar mixture of externalist and semantic holism leaves many questions unanswered, it is clear that it goes along with the externalist intuition insofar as the meaning of concepts is not taken by him to be determined by the 'internal' self-understanding (or the mental states) of a concept user. Rather, on Brandom's account, it is the *pattern of socially licensed inferences* between propositions involving certain concepts that determine their meaning (e.g. the fact that one can legitimately infer 'Fido is an animal' from 'Fido is a dog' determines part of the meaning of 'dog'). That inferences are socially approved of, however, is a matter of the *implicit social rules* of a linguistic practice as instituted in the entirety of a community's reactive dispositions (see Brandom 1994, chs. 1 and 2).

If we look from the perspective of this kind of externalist theory to the issue of the 'real' meaning of 'our' race and gender terms, we might say that the social conventions that regulate the use of such terms are essentialist insofar as they often support, for example, inferences from the ascription of racial or gender identities to ascriptions of intrinsic properties of certain kinds. In other words, our social conventions governing the *socially shared* linguistic use of race and gender concepts might only make sense on a metaphysically unacceptable essentialist interpretation of these terms. If we neither want to accept a naturalist analysis nor say that these concepts do not refer to anything at all, then it is very useful for a critical project to distinguish, once more, between the operative concepts at play, concepts that are determined by the objective social kinds that race and gender terms *really* track, and the manifest concepts which are determined by the *implicit, practical, collective* linguistic self-understanding of a community.

If we consequently locate 'manifest' concepts on the level of communal discursive practices (as would be completely consistent with Haslanger's

theory),[4] then it is plausible to say that *both* the manifest and the operative concepts are (at least somewhat) independent from *any individual's* self-understanding, and that both the existence of the manifest concept and the existence of the entities and properties tracked by the operative concepts are fully dependent on social practices. It then becomes possible to distinguish between *manifest concepts on the individual and the social level.*

Especially in the case of race and gender terms, we can imagine that there are contexts in which *some* competent users of these concepts do not explicitly believe, even on a close inspection of their semantic intuitions, that these concepts denote essential, intrinsic properties of people. For example, we might imagine that there are philosophers in some university who – as a matter of explicit belief – do not think that women or members of racialized groups are intrinsically less capable at doing academic work than men or 'white' people. Nevertheless, the very same individuals may still – without any sense of alienation – participate in a discursive practice in which inferences from ascriptions of race and gender to ascriptions of intrinsic philosophical abilities are regularly counted as valid and are treated as unproblematic. In a case like this, one could say that the *inferential norms* governing the use of their gender concepts in their discursive practice determine both their socially shared manifest concepts and their operative concepts, *despite* their explicit semantic beliefs.

Such a mismatch between the explicit self-understanding of individuals and their *collective self-understanding as instituted in a social practice* seems possible because discursive practices are constituted by more than just the individual semantic intuitions or the beliefs of their members. Socially shared inferential norms do not normally only reflect a sum of contingent individual mental states or dispositions, they also usually express practical distinctions that have belief-independent support in legal rules, conventional procedures, institutions, and material arrangements.[5]

If one acknowledges, however, that *both* the socially shared manifest and the operative concepts of a community are a matter of social practices which are not exhausted by discursive interaction in a very narrow sense but are highly interdependent with legal rules, organizational rules of

formal and informal work, with markets, city layouts and all other kinds of material arrangements, a narrow distinction between critical interventions that aim at a *better self-understanding* and political interventions that change the *institutional and social structure* becomes problematic. If we examine language as one component of social practices amongst others, a 'merely semantic' critique of race and gender naturalism that aims at an *improvement* of the self-understanding that a group has of its own conceptual schemas - that is, a critique that attempts to change the beliefs of group members about the meaning of concepts such that these beliefs better track the real meaning of these concepts - becomes a problematic idea if one separates this activity too much from social critique in a broader sense.

For example, if as a result of being convinced by a critique of gender naturalism someone only changes their individual understanding of the corresponding terms without ceasing to participate in the relevant discursive practice, then they would necessarily begin to use these terms *inadequately* according to the rules of this very practice. That is, a belief of theirs to the effect that they have now grasped the 'real' meaning of these concepts turns out to be false, at least in one sense: what they now take to be the correct rule of application for these concepts (especially in regard to inferences) is apt to generate failures of communication with their fellow language users. Thus, they must acknowledge that the critic has brought them to *change* their use of the concept rather than to *better understand* the collectively shared use.

This line of thought not only supports the conclusion that an immanent critique of race and gender that only aims at convincing people to *individually* revise their semantic self-understanding might miss its aim as long as it does not *also* aim at changing collective social practices that are *both material and discursive*. It also becomes clear, more importantly, that the 'real social kind' which our race and gender terms track is not metaphysically independent from the social practices that determine the socially shared manifest meaning of the relevant concepts. Rather, whenever we aim at revising our semantic self-understanding so that it better conforms to what race and gender 'really are', the very same process of revision might change exactly that, namely, what race and gender *really*

are. To put it more concretely, it might turn out that a society where it is collectively acknowledged, as a matter of inferential discursive practice, that 'our' concepts of race and gender track social kinds constituted by certain structures of oppression, could very well be (in virtue of the revisions to the whole web of practices that such a change would necessarily entail) a society in which these structures of oppression no longer exist in their present form.

To pursue this line of thought just one step further: if we acknowledge that certain social kinds (such as race and gender) can perhaps only exist in societies in which the inferential practices embody a *collective illusion* about the meaning of the terms by which these kinds are tracked, the goal of revising our semantic intuitions such that manifest and operative concepts become *congruent* does not make sense any more. Rather, what we then should ask is how our practices must change in order for there to be no longer any necessity for incongruence. This, of course, is a point that – in different ways – has been made both by Hegel and Marx. With his idea that there is a 'life of the notion', that is, a movement of continual conceptual revision that might never come to an end, Hegel has expressed the intuition that, at the foundation of our conceptual practices, there might be a social process that is not guided by a goal of eventual 'correspondence' between concepts and reality but rather by the goal of pursuing solutions to the specific problems of each stage in that process (see Hegel 1997, 27 f.). Against the potentially conservative implications of that theory, Marx (1940) has pointed out that it might be part of that very process that certain concepts of the social are 'necessarily misleading' inasmuch as their deceptive nature is rooted in the very social practices they seemingly enable us to understand. A further examination of these lines of inquiry might lead us away from questions of metaphysics towards a more sociological inquiry concerning the internal dynamics of social practices and historical changes, an inquiry that no longer relies as much on the criterion of whether our concepts correctly grasp objective kinds and focuses more on standards of progress that are internal to social practices.

3. The Critique of Ideology

The search for an immanent form of critique is, as I understand it, also at the core of Haslanger's notion of ideology critique. As the term is used throughout *Resisting Reality*, 'ideology' does not only include explicit beliefs, but also background assumptions, habits of thought and perception (18, 448) and socially shared schemas (413-418). These elements all form a background for the application of concepts (413). Ideology critique is thus thought to disrupt a *dogmatic* application of concepts (17), that is, an application of concepts that is insufficiently understood and insufficiently responsive towards considerations about what the *point* of using these concepts should be.

Haslanger argues that we should understand statements that employ ideologically-laden terms not as straightforwardly false, but as true in relation to certain *contexts of assessment* (419-421). In other words, ideological concepts (such as 'cool' or 'cute', in Haslanger's examples), if they are understood as relative to a social context of assessment, can be employed correctly without ideology critique becoming inappropriate – for such critique does not primarily rest on the assumption that the application of these concepts is *false*, but rather that these concepts (and consequently, the social contexts of assessment) are in some way *defective*. For this reason, ideology critique need not endorse the relativist implications that seem to follow from the idea of contextual truth. Ideology critique as a practice, Haslanger argues, must rather aim at finding a *common ground of assessment* from which a rejection of ideological judgements can be shown to be justified (425). Of course, such common ground cannot always be assumed to exist. Rather, ideology critique must often try to change or resist certain forms of *de facto* common grounds.

Even though Haslanger resists any temptation to enter into debates about normative justification, aspiring instead to 'only' elucidate the social reality that such debates are about, the notion of ideology is one in regard to which it becomes problematic to refrain from taking up a position on normative principles: for the question naturally arises as to how participants could ever be *justified* in privileging one form of common ground over another. Haslanger suggests that there might be formal criteria for designating some forms of common ground as superior, for example, if they can be reached without coercion or violence (426). These formal criteria are clearly useful, but it is not obvious that they capture everything that is wrong about ideologies, many of which precisely serve to support social structures of oppression without resort to violence. On an everyday understanding of 'ideology', ideologies are clearly in some sense *deficient* as *representations* of the world, but if they are, at the same time, true according to some contexts of assessment we should also expect the question as to which one of these contexts one should privilege to be answered in reference to certain normative *and* epistemic considerations.

Although I can only voice an intuition here that is not yet entirely developed, I believe that this problem is, once more, connected to the question of what the 'common' in common ground means. An agreement about certain conversational implicatures or an agreement about the applicability of certain concepts that constitute a common ground is, as Haslanger describes it, both a question of habit and socialization and a question of individually shared beliefs. What makes a 'common' ground into something that is *shared* can be understood, in regard to these aspects, as an agreement of *individual* dispositions or beliefs.

However, there might be another aspect of what makes a common ground into something 'common', namely a *normative understanding*, not only in regard to which further discursive moves are justified given that common ground, but which also might include an agreement about how one can *change* that common ground.

To use Haslanger's example: a conversational agreement about certain features of ethnic groups might be more or less problematic according to the degree the vocabulary of 'race' itself can be *challenged* within the further development of the conversation. We can compare two cases here: in the first case, there is an understanding that racial ascriptions (or judgements about 'coolness' or 'cuteness') might turn out to be false but that they can *never* turn out to be *inapplicable*. In the second case, however, while racial ascriptions are collectively accepted by all participants, they also (implicitly) accept the rule, that there are certain observations that would make it rational to drop this kind of ascription altogether from the

discursive repertoire. While the second case seems to be a case in which there are inappropriate or even misleading categories in play, it seems not such a clear-cut example of ideology as the first case. This is because, in the first case, there is a normative *restriction* built into a discursive situation which disallows certain ways of further developing the common ground which does not exist in the second case. Such a restriction on the ways in which a common ground can be changed (a restriction which is especially obvious in the case of the ideologies of race and gender naturalism) is a *normative* feature of the discursive situation which – while it might be supported by habits or beliefs – requires more than a mere *agreement in content* between the respective interlocutors' independently existing beliefs or dispositions. This normative feature rather belongs to the institutional aspects of a speech situation, or rather, of the normative social practice into which the 'common ground' is embedded.

If we acknowledge this normative aspect of ideological 'common grounds', we might formulate an immanent critique of the degree to which the institutional or normative rules of conventional speech situations *limit* the development of conceptual alternatives from 'within'. This might turn out to be a criterion to distinguish ideological from non-ideological 'common grounds' that allows for degrees, and that combines normative and epistemic considerations in just the right way (see also Stahl 2013). Of course, as soon as we understand this to be not only a matter of individual dispositions and beliefs but also of material practices and forces, the 'common ground' might turn out to have a richer material and historic dimension than those which become accessible if one only focuses on speech-act theoretic considerations alone.

Titus Stahl is Assistant Professor of Philosophy at Goethe University Frankfurt and a co-editor of the Zeitschrift für philosophische Literatur. He has published two books: *Immanente Kritik Elemente einer Theorie sozialer Praktiken* (Frankfurt/M.: Campus 2013) and *Einführung in die Metaethik* (Ditzingen: Reclam 2013).

References

Brandom, R. B. (1994) *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, MA: Harvard University Press.

Burge, T. (1979) 'Individualism and the Mental'. In *Midwest Studies in Philosophy* 4: 73-121.

Haslanger, S. (2012) *Resisting Reality. Social Construction and Social Critique*. Oxford: Oxford University Press.

Hegel, G. W. F. (1977) *The Phenomenology of Spirit*. Translated by A. V. Miller. Oxford: Oxford University Press.

Lepold, K. (2013) 'Sally Haslanger: Resisting Reality'. In *Zeitschrift für Philosophische Literatur* 1(1): 26-34.

Marx, K. (1940) *The German Ideology*. London: Lawrence and Wishart.

Marx, K. (1972) 'Contribution to the Critique of Hegel's Philosophy of Right. Introduction'. In *Critique of Hegel's Philosophy of Right*. Translated by A. Jolin and J. O'Malley. Cambridge: Cambridge University Press, 129-142.

Putnam, H. (1973) 'Meaning and reference'. In *The Journal of Philosophy* 70, 699-711.

Saul, J. (2006) 'Gender and Race'. In *Aristotelian Society Supplementary Volume* 80(1), 119-143.

Stahl, T. (2013) 'Ideologiekritik als Kritik sozialer Praktiken. Eine expressivistische Rekonstruktion der Kritik falschen Bewusstseins'. In R. Jaeggi, D. Loick (eds.) *Nach Marx – Philosophie, Kritik, Praxis*. Berlin: Suhrkamp, 228-254.

---

[1] Where not otherwise indicated, all references in brackets are to Haslanger 2012.

[2] This variety of externalism was introduced by Putnam (1975). Haslanger's full definition is 'natural kind terms or 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' (374).

[3] The most famous proponent of this variety is Burge 1979; Haslanger characterizes it as follows: 'the meaning of a term or the content of a concept used by a speaker is determined at least in part by the standard linguistic usage in his or her community.' (374)

[4] In her discussion of the 'improvisation theory of meaning' (437-439), Haslanger also requires a 'coordinating intention' of language users in respect to the usage in their community as well as a 'shared tradition', but this still allows for the question of whether that which makes up these shared elements is, in any way, independent from individual semantic beliefs and intuitions (see also Lepold 2013, 33).

[5] Carefully distinguishing between individual and shared manifest meanings might also be helpful for countering objections such as those by Saul (2006), who argues that a survey of uses of race and gender terms rather supports that they do not have any clear meaning at all. Against this view, it could be argued that while it might be true for many individuals, there can still be social rules that give these terms a (relatively) clear meaning.