

Categories We Live By: Reply to Alcoff, Butler, and Roth

Book symposium on *Categories We Live By: The Construction of Sex, Gender, Race, and Other Social Categories* by Ásta. Oxford University Press, 2018.

Forthcoming in *European Journal of Philosophy*.

Ásta, Duke University

<https://orcid.org/0000-0003-0252-6527>

Abstract

The author of *Categories We Live By* replies to critics Linda Martín Alcoff, Judith Butler, and Abraham Sesshu Roth.

Introduction¹

There is a tremendous amount of philosophical interest in social phenomena today, even in theoretical philosophy. When we look at various current research approaches in epistemology, or philosophy of science, or language, or metaphysics, for example, social aspects are prominent. This is a real shift from how philosophy was practiced at the end of last century. But there is also an increased interest in theorizing the social world itself.

For the earlier generation of philosophers working on social ontology, theorizing the social meant theorizing the institutional. Theorizing the institutional is about agreement or common plans or joint intention. And it is about how we can get collective agency from us individual little agents.

When the paradigm of the social is the law or institutions like universities and corporations, social categories like gender and race are an afterthought. Can collective acceptance really be what keeps these categories in place? The tools that were fashioned to explain how one might have a desire-independent reason for acting in a way that isn't moral seem ill equipped to deal with the messiness of categories that no one owns up to wanting to keep around.

In *Categories We Live By* I take a different tack (Ásta 2018). The paradigm of the social that is to be theorized is not institutional categories, although the framework I developed can be used for both institutional and non institutional, or what I call "communal" categories of individuals. These are categories we find ourselves in, thrown into, if you will, often even against our will.

The motivation for theorizing social categories of that sort is not to simply describe reality or explain how the status quo is maintained. I have my doubts

about the possibility of even engaging in that sort of project. No, the project is self-consciously guided by a commitment to feminist intersectional liberatory politics.

The aim is to offer a theory of all social categories of *individuals*. This means that I don't intend to offer a theory of other social categories, although I think the framework can capture some other social categories besides those of individuals.

The scope is broad—perhaps too broad, if Linda Alcoff is right (see essay in this volume)—all social categories of individuals, even though gender is a paradigm for me and the initial target.

A picture of social categories

Let me offer you the big picture. Think of categories like women, men, and genderqueers, for example. What is it to belong to one of these?

There is a tradition in philosophy, and in the various cultures we are all part of, which has it that you are a woman, say, entirely because of something about you. There are features of you that make you a woman. A lot of feminist theory and philosophy has been focused on figuring out what that is. For after all, we need to know who the women are if we are to end their oppression. And we get into trouble because each attempt at a definition runs into problems. We privilege one feature over another and that is bad theoretically, and politically, and morally. Are you a woman because you are nurturing and caring? Are you a woman because you are perceived or imagined to have female reproductive organs? Are you a woman because you identify with a woman's social role or with feminine norms and take them as applying to you? There must be something about you that makes you a woman, darn it!

So here is an idea: for the most part, it ain't about you.

What do I mean? I mean that what makes you a woman or a man or genderqueer or what have you is mostly about other people. It is about other individual people, to some extent, but mostly about the societal setting you find yourself in, with its values, norms, practices, and, yes, material conditions. Those are the things that determine what is possible for you to be and other individuals and groups assign you roles in accordance with that. They are the foot soldiers who enforce the values of the social setting you find yourself in. The foot soldiers who enforce the ideology, if you will.

We can cast what I am suggesting in different ways. My preferred way is to say that gender and other social categories are *deeply social*.

On this deeply social picture you are a woman, not because of something about you, but because of something about other people and the society at large: it is their interests and values that place you in a category. You are a woman because

other people care about features that they take you to have. And to be a woman is simply to have a certain social status in a context.

The deeply social picture involves a reorientation of the sort I take Judith Butler to have been advocating since *Gender Trouble*, although the details will differ (Butler 1990, 1993).

This deeply social picture is most notably contrasted with the post-Beauvoirean feminist picture, still immensely popular today, where sex categories are natural and given and gender categories are the social significance of sex. Gender, on that picture, is the way society values sex. You are of a certain gender because you have, or are taken to have, certain sex characteristics and the society values those in particular ways. On that picture, you are of a certain gender because of something about you (your sex or your perceived sex) but the meaning and consequences of that depend on the societal values.

On the other hand, when we think of these categories as *deeply* social, then societal values dictate which features of individuals are important and people get conferred onto them social statuses in accordance with that. The deeply social picture opens up yet other ways of taking a critical stance towards the world we live in. On this picture, society may value certain bodily features over others and people get conferred statuses in accordance with that. We can question that these features are valued in the way that they are, we can question the content of the conferred roles that come with being taken to have those features, and we can question the associated norms for playing those roles well. This is not to say that there could not be a good justification for valuing certain features in certain contexts. In many contexts there may be reasons to value capacity to bear children, for example, or being exceptionally tall, or good at programming. But what the deeply social picture offers us is the opportunity to ask, not only for the explanation for why something is valued, but also its justification. And in many contexts the justification is missing or spurious.

So the suggestion at the most basic level is this: it is a reorientation. There are a lot of nuances, qualifications, and complications, but the basic reorientation will remain: ask not only what about you makes you a woman, ask what about society does.

How visible is this big picture in the book itself? Perhaps not very. Many people will read the book as a description of the work of the foot soldiers, who individually or jointly keep conferring social statuses on each other. This picture just sketched goes some way towards approaching the question why the features that are socially salient in a context are socially salient. I am mostly silent on that question in the book. I also think people could use alternative explanations for the social salience than my preferred one. But I hope this big picture will help in coming towards Linda Alcoff's concerns in particular, so let me start with them.

Linda Alcoff

Agency and the threat of the static

Linda Alcoff raises three important worries about my approach. One is skepticism about attempting to give a theory of all social categories of individuals. Another concerns worries about what happens to human agency on this theory: am I not giving us all too much agency in the construction of these categories? And the third one is a concern that while I want to offer a conception of social meaning, and what it is for a feature of us to matter socially, the resulting account is too narrow and that there are ways in which features of us matter socially that are not captured by my account.

I think the big picture I sketched before goes some way towards meeting Alcoff's worry about agency. Why are the features that are socially salient socially salient? That is because they are valued. By whom? To say that they reflected the values and interests of the participants would be too simplistic. We all enforce hierarchies and arrangements that we may not value, but we are part of a community that values them, or are subject to the enforcement of those values. The general social values inform the underlying maps we bring with us to each context. They reflect societal hierarchies and values. In a society infused with racism and sexism, for example, the societal values will reflect that, as will the social maps people bring to each encounter. Even if the participants are themselves anti-racists, the social maps they bring to the encounter will reflect the values of the larger culture, inflected by the stance of the participants towards the larger societal values, and infused with their own values.

There is a recurring worry about the picture that I offer in the book that it is too static and ahistorical. I think that is a problem and the metaphor of the social map perpetuates that. I hope that it is not inherent to the theory, and that the theory can be expanded upon in a way that makes it truly dynamic. There are resources in philosophy of language that I think can help with that task.

What I have in mind are ways of making sense of intelligibility and permissibility conditions in discursive contexts in a way that is dynamic. I think adopting these ideas and adapting them for action contexts can help, although some worries remain, as we will see.

Consider Paul Grice's notion of common ground, as developed by Robert Stalnaker (Grice 1981, Stalnaker 1999, 2002). On Stalnaker's view, the common ground in a conversational context is a set of propositions or beliefs that are taken for granted by the participants in the conversation. Each conversational move changes that set as propositions are added or subtracted.² For this to be useful for our purposes it is important that none of this need be conscious. Making use of an analogous idea of a common ground for action contexts, as opposed to conversational contexts, we can then say that by placing a person on a

social map the subjects also place certain assumptions in the common ground, i.e., the set of shared assumptions in the context.

The placing of people on a social map is not always a simple affair that happens without struggle. Often the people in the encounter bring incompatible social maps and some negotiation happens before people settle into their roles. Some contexts are even too short for anyone to settle into any role. There are attempts at placing each other onto social maps, but it is contested from start to finish. The unfortunate aspect of the metaphor of a social map is that it suggests that we apply a fixed map that remains static through the encounter, but the suggestion here is that each action move in the context changes the shared assumptions, or the common ground a little bit. The social map is then a dynamic entity, changing ever so slightly with each move in the context. In each and every context we travel there are certain features that are socially salient and people taken to have those features get conferred onto them an intersectional status consisting in constraints and enablements on their action because being placed in a category is to have certain assumptions made about one that set the intelligibility and permissibility conditions for their actions.

Although I believe we can in that way make the context of the encounter *dynamic* and account for the historical dimension in that way, where the frame of intelligibility is an ever changing beast, the model has other problems. In particular it appears too cognitive. Most of us are not aware of what assumptions, including social and cultural, are brought to the encounter. We are, for example, not aware of how our gestures and body language contribute to placing and keeping each other in our social roles in the encounter. Those are part of the placing and enforcing mechanisms just as much as various speech acts are. But can we make sense of the idea that we place assumptions in a common ground where most of the time these are unconscious, and where moreover, gestures and body language can be moves that are the equivalent to conversational moves in the discursive context? If we can make sense of that, then the model starts to look more palatable and promising. Obviously this all has to be worked through in some detail.

Social meaning

This brings me to Alcoff's second worry which concerns my intention to capture what matters socially with this framework. I think Alcoff is right to worry about this and a similar concern has been raised by other critics, including Elisabeth Barnes and Matthew Andler, Rebecca Mason, and Åsa Burman (Barnes and Andler 2019; Burman 2019; Mason 2020)

The conception of social meaning that I offer in the book is this: a feature has social meaning in a context just in case people treat you differently if they take you to have it. The precise formulation is:

a feature **B** has social meaning (significance) in a context in which individuals taken to have **B** (they need not actually have it) have conferred upon

them another feature F, which is a social status and consist in constraints and enablements on their behavior in the context.

Have I captured everything one might want to capture with the notion of *social meaning* or *social significance*? No, I have not. I have offered a certain conception of social meaning. A feature can be “socially meaningful” in a way that this conception does not capture. For example, consider my grandmother. She got polio at the age of two and had certain difficulties with walking and so on as a result. Also, if she slipped and fell on the floor, she could not get up except by crawling to a place where she could lift herself up with her arms. She was deeply ashamed of her physical impairment and hid it and refused to think of herself as disabled. People who met her had no idea and did not treat her differently from people without physical impairment. But her physical impairment had an effect on her socially. She would avoid events that included a lot of walking, especially in uneven terrain, and at a party might not circulate, but stand somewhere where she could grab onto things. Over time this led to various changes in people’s behavior towards her and hers towards others, in a way we might say has social significance. It is, however, not captured by my conception, nor is it intended to. I would rather describe this as social consequences of her having the physical impairment, and involve a different conception of social construction to understand the case. I discuss several different conceptions of such social construction in Chapter 2 in *Categories*.

Why attempt to give a metaphysics of all social categories of individuals?

I now come to Alcoff’s most basic worry. Why attempt to give a metaphysics of all social categories of individuals?

It is of course with this sort of offering that the proof is in the pudding. If it is useful, it will have been worth doing. If it isn’t useful, it will have been effort in vain. If it can help illuminate aspects of the social world that need changing and give us ideas for how to change them, I will be happy. But I am aware that each attempt at illumination will bring some things to light and eclipse others. That is why we have to keep front and center what we want the theory for. There can be other attempts at illumination that have a different purpose.

This is a picture of that part of social reality that is constructed from our valuing certain features of individuals over others. It gives us a chance to question that valuing for each and every context, along various dimensions. But that sort of questioning will require us to go into the messy details of each case, its history and function, in a way that this somewhat skeletal presentation of the theory avoids.

I am glad that Alcoff pushed me to the expand on what I say in the book to gesture at what a fuller story would look like. Let us consider Alcoff’s example of race at the end of her comments. The resources in the book to illuminate that case involve both the mechanisms of institutional statuses, and the mechanisms

of the communal statuses. I agree that if we got rid of the institutional racial statuses, we would still have racial hierarchies. The fuller story of what happens at the communal level would involve the claim that hierarchies are enacted and enforced communally in the contexts we find ourselves in, by us foot soldiers of racial ideology, often unbeknownst to us, even if we are committed anti-racists. The mechanisms at the communal level start to look a lot like ideological apparatuses and enforcing mechanisms of the Althusserian sort (Althusser 1970).

Judith Butler

Let me now turn to Judith Butler's criticism. There are a lot of things here. I appreciate the difficulty of attempting to reconstruct one's thinking from decades ago. The funny thing is that the first version of that chapter was written some time ago, too, in Sally Haslanger's graduate seminar at MIT in the late nineties. But I stand behind that interpretation, all the while acknowledging the problems Butler raises.

I want to speak to three things here. Two of them concern the interpretation of Butler's early work, but they also link to how to develop the ideas in *Categories*. The third is the question of methodological individualism.

The use of Austin's work

Let me start with the use of Austin's work (Austin 1975). I agree that one has to modify Austin's work considerably to use it to make sense of some of these phenomena. In making sense of social categories of individuals I wanted a theoretical tool that could capture both that people could be attempting to track some independent fact, but in that attempt they also assign a status, and that it is the status that ultimately matters. Austin himself does not offer this. He is, of course, interested in the use of certain verbs, as opposed to types of action, but let us put that aside. But the main issue is that he has these two types, exercitives, where you are exercising a power, and verdictives, where you are attempting to track a fact and render a verdict as to that fact. Austin's exercitives are divorced from the activity of attempting to track a fact, and the verdictives are oblivious to what happens when you issue your verdict, to the fact that when you issue your verdict you are classifying or placing the individual and then that classification has a life of its own. The conferralist framework is meant to capture these. But it is also broader in scope as it concerns all actions, not just discursive actions. For those of you who are wondering, using Searle's taxonomy of speech acts does not resolve the problem (Searle 1969).

Returning to the use of Austin, I take it that another infelicity in using Austin's framework for Butler's purposes is that it suggests a person is given a status once and for all in a public explicit act, but, quite on the contrary, the notion of iteration and of sedimentation play a large and important theoretical role for Butler. So some modification of the Austinean picture would have to be made.

I like Butler's example of the performance of same sex marriage on the steps of city hall in San Francisco before marriage equality. These are ways of laying a claim to a legal status. But are they more than exhortative acts, just like mock trials, or a king's beheading on a stage? I think they play a role in the ongoing discussion of what the basis for marriage should be. What should the entry requirements be? Or as I would put it, what should the base properties for the conferral of the status be? People are questioning the then current base property for the conferral of the status and enacting a scene where the base property is more just.

The game analogy

Let us turn to the game analogy. I agree that there are big problems with it. And I think it is instructive to look at its failure.

One thing the analogy with the game is supposed to help us see is how only certain phenomena assume meaning within a certain activity. If we take baseball, for example, there are a lot of bodily movements that happen on the field, but only some of those count for the purposes of the game. Someone scratches their ear or moves the bat in a certain way as they wait on deck, and none of that matters. That very aspect of a game is still helpful, I think. But then we go on. What determines which bodily movements matter are the *rules* of the game. They determine which body parts and which bodily movements have meaning within the game. What is unhelpful about the analogy, I take it, is that, unlike baseball, gender does not have explicit rules and whatever implicit rules or norms there are is constantly changing.

It seems to me that some of the shortcomings of the game analogy can be overcome. For example, roles can be assigned before birth. We don't choose these roles, but are thrown into them. It isn't a voluntarist picture at all. But if it isn't a voluntarist picture, why is it behaviorist? I am not seeing clearly how that need be. I agree that the fuller picture needs to leave room for the role of imagination as well as the imaginary and involve a story of the psychosocial development of the subject that is robust. But I see no reason to think that the only picture of human psychology compatible with this story is one that has us be reactive systems responding to stimuli. I agree that a robust story of psychosocial development is absent in the book, but I hope that such a story is compatible with what is in there.

Another problem with the game analogy was that it suggests that the activity in question is governed by rules that are fixed. Earlier I sketched a way to make the schemas of intelligibility a dynamic set that was ever changing. In that way, it can also have a history. I don't know to what extent that is compatible with Butler's picture. It seems to me that to make sense of conditions both of intelligibility and of permissibility it isn't enough to only talk of norms. Norms help us make sense of what it is to perform well or badly but cannot explain why they apply to us at all. Why do these sets of norms apply to us rather than those? We

cannot flout a norm that doesn't apply to us, but what determines whether a norm applies or not?

On my own story, if you are taken to have the base property for the conferral of a status, then the norms apply to you, whether you like it or not. This can, of course, be deeply unjust.

Methodological individualism

That brings us to the last topic, methodological individualism. That is a recurring theme, and has been raised by other critics. Am I right to think that Butler and I agree about what is to be theorized? If so, then we agree on what the subject matter is, but I think that I can account for it with a more sparsely populated set of tools. In particular, I don't, at present see the need to posit the existence of structures, or plural wills, or ontologically basic plural subjects. I think there are structures, but they are the result of the activities of individual agents. But my commitment to methodological individualism is not religious. I will give it up if I need to. It may be that the tools I have cannot do the work. I will then adopt or even smith some new ones. One of the things that still stands out to me as a weakness of the theory offered is how to approach the role of materiality and historical memory. The conception of ideology that can accompany this theory can incorporate some material factors, but this is and remains a story that makes valuing some features over others the central idea and is, in that regard, privileging the ideal over the material.

Abraham Roth

Let me turn to Abraham Roth's comments. Roth's main concern is whether the conferralist framework can really account for the stability of social categories of individuals. If it cannot then it is of limited use for social research.

A lot of people working in social ontology, who are interested in social categories, approach them from a slightly different angle than I do. They are interested in the possibility of social kinds, where a kind is understood as a grouping of individuals or stuff that has explanatory and predictive power and sometimes even shares something like an essence. Natural kinds are the paradigm, where a natural kind is understood as a grouping where the explanatory and predictive power is due to the essence of the kind. For example, we can predict and explain the behavior of gold under various conditions from knowing the essential chemical features of gold. This is the conception inspired by Aristotle, but there are others, of course, including the one that has been applied to human kinds recently by Ron Mallon, which relies on Richard Boyd's idea of natural kinds as homeostatic property clusters. (Boyd 1999; Mallon 2016)

Are there social kinds in that roughly Aristotelian sense? Take women, for example. Can we isolate a feature or set of features that all and only women have

that explains the behavior of women? You all know the answer to that, no we can't.

But wait a minute. What if we give up on the idea of a substantial answer? What if we say that the defining feature of being a woman is just to have a certain status in a context?

But then there are at least two problems: which status are we talking about? And what about intersectionality?

This is where the base properties come in. You have a status in a context on the basis of being taken to have a feature that is socially salient in the context. If there are more than one property socially salient, the status you have reflects the perception that you have a combination of those features. The status is thus truly intersectional.

So, even though I don't approach the question of social categories from the social kinds angle, there is a way to link my project to those projects, and there is a way to ask and answer some of the questions such a theorist may be interested in.

But then we get to the main worry Roth has, which concerns the stability of social categories, given the framework I offer.

I take it that the worry is that social categories, especially if thought of as something like social kinds, are assumed to be stable across contexts, but that it is hard to see how the framework I offer can explain that stability.

We need such stability to do analyses in social research, as well as ground claims of solidarity across contexts, contemporary and historical. For example you may have seen the new Global Gender Gap Report issued by the World Economic Forum in March 2021 (World Economic Forum 2021), where the not so surprising finding was that the gender gap has increased under Covid-19, presumably explained by the fact that all the care work that has been happening in the last year, mostly unpaid, has been done predominantly by women, often on top of other work. But in order to engage in that sort of research and get that sort of result we need a way to justify the claim that an individual in India and in the US and in Iceland are all the same in one respect, they are all women. Can my theory, as radically contextual as it is, ground that claim?

Roth is right that the base properties link statuses in different contexts to each other. We can lasso the individuals in the different countries into one big set by focusing on all the people who have a social status on the basis of being taken to have some certain base properties. The base properties in question do not need to be just sex assignment, but can also include other base properties that have historically been associated with a certain sex assignment. So here how we get all the women together: the people who have a status on the basis of being taken to have the sex assignment female or any of the properties historically associated with the feature of being female, including having feminine gender norms

apply to you, allying yourself with the social status of women, and so on. Then the question is: how have these individuals fared in the Covid-19 pandemic? And the answer is that they are lagging further behind economically and in other ways than the last time such a study was made.

I think Roth is right that there is a real concern about how stable these categories are. I think I have some resources to address that worry, in the way I have sketched above. But it may not be enough. I welcome the augmentation of the picture that involves co-conferral. But is that best thought of as pertaining to social categories or is it about the interaction of other relationships that people stand in, for example friendship or romantic partnership and their membership in social categories?

I want to speak briefly to the question how to make sense of the couple who was legally married for decades, but who shouldn't have been. I want to offer an alternative to Roth's suggestion regarding that case. I still maintain that they had the legal status throughout that time, with the associated legal constraints and enablements, joint filing of taxes and a host of other things. And apart from that legal status they also had an intimate relationship and the practice of that relationship was the source of other duties they had towards each other. For these two individuals stand in many different relationships with each other, and the many relationships are the source of many different duties, —some of which can even pull in different directions at times, as we know from our own experience, as well as from epic poetry and drama.

References

Althusser, Louis. 1971. "Ideology and ideological state apparatuses (Notes towards an investigation)." In *Lenin and Philosophy and Other Essays*, pp. 142–7, 166–76. Translated by Ben Brewster. New York and London: Monthly Review Press, 1971.

Ásta. *Categories We Live By*. New York: Oxford University Press, 2018.

Austin, J.L. 1975. *How to Do Things with Words* (2nd ed. Ed by J.O.Urmson and Marina Sbisa). Cambridge, MA: Harvard University Press.

Elizabeth Barnes, Matthew Andler, Review of *Categories We Live by: The Construction of Sex, Gender, Race, and Other Social Categories*, by Ásta, *Mind*, Volume 129, Issue 515, July 2020, Pages 939–947, <https://doi.org/10.1093/mind/fzz041>

Burman, Ása. 2019. Review of Ásta *Categories We Live By: The Construction of Sex, Gender, Race, and Other Social Categories*, Oxford University Press, 2018. *Notre Dame Philosophical Reviews*. 2019.06.05

Butler, Judith. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. New York and London: Routledge.

- . 1993. *Bodies that Matter*. New York and London: Routledge.
- . 1993. “Critically Queer”. *GLQ: A Journal of Lesbian and Gay Studies*, November 1993.
- . 1997. *Excitable Speech: A Politics of the Performative*. Routledge.
- Grice, H. Paul. 1981. Presupposition and conversational implicature. In *Radical Pragmatics*, ed. by Peter Cole. New York: Academic Press. Pp. 183-98. Reprinted in H. Paul Grice *Studies in the Way of Words*. Cambridge MA: Harvard University Press. 1989, pp. 269-282.
- Lewis, David. 1979. “Scorekeeping in a language game.” *Journal of Philosophical Logic* 8 (1):339--359.
- Mason, Rebecca. 2020. Review of Ásta. *Categories We Live By*. New York: Oxford University Press, 2018. *Hypatia Reviews Online*, 2020.
- Mallon, Ron. 2016. *The Construction of Human Kinds*. Oxford University Press.
- McGowan, Mary Kate. 2019. *Just Words: On Speech and Hidden Harm*. Oxford University Press.
- Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. London: Cambridge University Press.
- Stalnaker, Robert C. 2002. Common ground. *Linguistics and Philosophy* 25: 701-21.
- Stalnaker, Robert. 1999. *Context and Content*. Oxford: Clarendon Press, 1999.
- Tirrel, Lynne. 2012. “Genocidal Language Games.” In *Speech and Harm: Controversies Over Free Speech*, Ishani Maitra and Mary Kate McGowan, eds., Oxford: Oxford University Press, pp.174-221.
- Tirrell, Lynne. 2018. “Authority and Gender: Flipping the F-Switch”. *Feminist Philosophy Quarterly* 4 (3). <https://doi.org/10.5206/fpq/2018.3.5772>.
- World Economic Forum, Global Gender Gap Report 2021. Geneva, Switzerland, 2021.

¹Acknowledgements: I would like to thank Linda Alcoff, Judith Butler, and Abraham Roth for their incredibly generous, and challenging, criticism of my work. It is an honor and privilege to have the opportunity to engage with them about these issues.

I would also like to thank people who helped put the APA session together, where this symposium took place, originally scheduled for 2020: Rebecca Chan, Rebecca Mason, Katharine Jenkins, Charlotte Witt, and Nico Orlandi. Last, but not least, I want to thank my partner, Dore Bowen Solomon, for making this book symposium possible, during lockdown and school closures. Of course, none of the people mentioned are responsible for the views expressed in *Categories* or in these pages.

²David Lewis’ idea of keeping score in a language game can be useful here too (Lewis 1979). See also

further developments by Mary Kate McGowan (2019) and Lynne Tirrell. (Tirrel 2012, 2018)