# BOOK REVIEW: MARQUES, T. & WIKFORSS, Å (EDS.), *Shifting Concepts* (Oxford University Press, 2020, 284 Pages).

## JOAN GIMENO-SIMÓ

https://orcid.org/0000-0001-9771-9385
*University of Valencia*
*Department of Philosophy*
*Valencia*
*Spain*
joan.gimeno@uv.es

**Abstract**: In this review I provide a brief analysis of the main features of the collective volume Shifting Concepts (Oxford University Press, 2020), edited by Teresa Marques and Åsa Wikforss. The volume addresses several related topics, and it contains contributions from psychologists and philosophers. It deals with the topic of concept variation understood in a broad sense, for it tackles diachronic, contextual, interpersonal and even intrapersonal variation; besides, the second part of the book is devoted to the topic of concept revision and amelioration. I provide a brief description of the book and then I critically assess each of the contributions.

*Shifting Concepts* (Marques & Wikforss 2020) is a collection of essays addressing the issue of conceptual variation from a wide variety of perspectives, both psychological and philosophical. It covers a vast range of topics related to the notion of *concept*, ranging from the current debate concerning the limits of conceptual engineering to questions pertaining to core topics in philosophy of language and mind, such as the dispute between externalism and internalism and the very nature of communication. The volume tackles conceptual variation in a broad sense, as it addresses contextual, diachronic, interpersonal and even intrapersonal variation, as well as intentionally induced change. Many chapters shed light on these issues by means of empirical studies, whereas others do so in theoretical terms. The wide variety of topics, as well as the large number of different perspectives from which they are addressed, make this volume a very comprehensive contribution to understanding the notion of *concept* and its current role in several areas of knowledge.

The volume is divided in two parts. The first part is devoted to conceptual variation in all the senses described above, as well as to the mechanisms that enable speakers to overcome such disparity in order to communicate successfully. Many of its chapters revolve around one of the classic topics in intercultural concept variation, that of *color concepts*. The second part, in turn, addresses several issues related to the topic of conceptual engineering: whether intentionally induced change is possible, how it can be attained and under what conditions it can constitute an amelioration of the concept. The concept of *race* is the most widely discussed example in these chapters. The whole book is preceded by an introductory chapter in which the editors present and summarize each of the contributions. Let us take a closer view to each chapter of the book.

The first chapter, by Barbara Malt, addresses an important question: the mapping between the lexicon and the kinds of entities in the world. The whole chapter is a sustained attack on the traditional picture according to which objects come in natural groupings each of which is assigned a word. Instead, Malt argues that the groupings tracked by naming patterns are often determined by properties that have nothing to do with naturalness. Phones are one such case. Modern smartphones bear far more resemblance to computers than they do to the first devices that went under the name "phones", and yet they inherited the label of their ancestors. This means that such lexical classification is not grounded on similarity, but rather on the history of the devices. Malt provides an array of similar examples that further undermine the traditional picture: cross-linguistic discrepancies regarding the range of application of equivalent words, naming patterns grown out of marketing (such as the choice not to classify tablets, in contrast with laptops, as a type of computer), etc. She also explores some consequences that the new picture has for issues like communication and language learning. Overall, the chapter provides powerful reasons to cast serious doubts on ideas that have dominated the philosophical literature for decades – chiefly, David Lewis's (1984) notion of *reference magnetism*.

The second chapter, by Gregory Murphy, also challenges one of the most prevalent philosophical doctrines: externalism. Bearing on evidence about the differences between children and adults's concepts, and relying on the fact that these disparities do not prevent communication, Murphy argues that sharing a common, public concept is far from being a necessary feature for ordinary exchanges. Rather, he claims, successful communication often comes in degrees, and all that it requires is that individual concepts overlap significantly. He dismisses public concepts as unnecessary for theoretical purposes.

Malt's chapter shed light on how conceptual differences across speakers and languages arise, whereas Murphy's contribution provided powerful arguments for acknowledging that communication does not require identity of concepts. The next chapter, by Peter Pagin, is in line with the former two, since it provides a useful framework for specifying both the circumstances under which conceptual differences can come to light and the several ways in which communication may succeed in spite of these differences. He begins by arguing that, in everyday exchanges, interpersonal differences in concepts often go unnoticed, for they are obscured by pragmatic issues. Consider two speakers who disagree on whether the concept of *arm* includes the hand. If the speaker with the inclusive concept is feeling a pain in his hand, he may truly report such feeling by uttering (1):

(1) I have a pain in my arm.

However, and here is the main point, such report would be highly misleading, given Grice's Maxim of Quantity (Grice 1975) and given the availability of a report like (2):

(2) I have a pain in my hand.

It is for this reason that even speakers with an inclusive concept of *arm* will ordinarily prefer to report their pain in the hand by means of (2). The corollary is that conceptual differences will remain hidden in a large part of everyday conversations. Pagin goes on to describe several features of the kinds of contexts that do make it possible to attest a difference – say, because implicatures are blocked. He also uses modal embedding for showing how conceptual differences can be brought to light. On the basis of these examples, he establishes a distinction between different

levels of communicative success: at the purely semantic level, at the level of what is said and at a third level in which implicatures are also involved. All in all, I believe that Pagin's contribution shows that the resources that semanticists and philosophers of language have been employing for decades can still be very useful even for approaches that acknowledge the challenges that Malt or Murphy put forward. It also provides interesting methodological suggestions that may help future researchers dealing with those very challenges.

The fourth chapter delves even further into the sloppiness of concepts. James Hampton reviews a series of empirical studies whose result is a negative one: none of them shows that intension and extension are strongly related. Namely, they show that the features that people typically associate with a concept often fail to determine the set of individuals that they would classify as falling under that concept. At least *prima facie*, these findings undermine another assumption that has dominated the philosophical tradition arguably since the work of Frege (1892): the idea that intension is able to uniquely determine extension. A response in externalist terms could hold that perhaps what these studies show is not that intension fails to determine extension, but rather that ordinary speakers are usually wrong about the extension or intension associated with a common, public concept. Yet the argument could be reversed: the mismatch deployed by ordinary speakers can be used for arguing that the notion of public concepts (i.e., entities able to determine both intension and extension) is helpless when it comes to the scientific task of predicting and explaining the way people categorize reality.

The next two chapters address a common topic: color concepts. In chapter 5, Yasmina Jraissati argues that (lexical) color categories, far from being fixed, are always sensitive to the context and purpose of the conversation. A wallet that would be labeled "brown" in most contexts can also be

referred to as "red" in case we intend to distinguish it from other similar wallets of a darker shade. We therefore need not just a notion of similarity, but also one of dissimilarity: objects are classified as belonging to one color category or another depending not just on how similar they are to other objects, but also in relation to the objects from which we intend to distinguish them. Thus, labels like "red" or "white" receive different extensions depending on their domain of application; when applied to wines, for example, they are employed to denote highly and poorly colored substances, respectively. Jraissati's discussion appears to be carried out in semantic terms (i.e., as if the context-dependence she describes were part of the very meaning of the words), although all her claims are, I believe, compatible with treating these phenomena as belonging to pragmatics – say, to treat the application of one label or another as a special case of Grice's Maxim of Relevance. At any rate, she puts forward many important issues that should be taken into account in future inquiries into the semantics and pragmatics of color terms.

Chapter 6, in turn, deepens still further into the instability of the extension of color terms. Its authors, Zed Adams and Nat Hansen, argue against two antagonic positions: those who claim that colors are intrinsic properties of objects and those who maintain that they are relational properties. Against these two rival views, Adams and Hansen claim that whether colors are perceived as intrinsic or relational depends on the very object we are dealing with. Thus, ordinary people may perceive color as an intrinsic property of, say, a table, but things get more complex if we consider mirrors or even rainbows. In view of the high amount of intrapersonal, contextual, and interpersonal variability displayed by color terms, the authors conclude that there is no "common-sense" notion of color, and that it is a historical development. The whole chapter may be seen as

an advice for philosophers not to rely on allegedly ordinary intuitions about the phenomena they are dealing with. Taken together, this chapter and the former describe a considerable amount of challenging data that will have to be tackled by anyone interested in describing how color terms work in natural language.

The topic of externalism is reprised in the last chapter of the first part of the book. There, Daniel Cohnitz and Jussi Haukioja try to establish a connection between the mainstream philosophical view of meaning and the intuitions of ordinary speakers, and they do so by arguing for an *externalist view of meaning* while at the same time maintaining an *internalist metasemantics*. Essentially, they argue that, even though the reference of natural kind terms like "water" is determined by external factors, what ultimately settles whether a particular word is being used as a natural kind term or as a functional kind term are the intentions of the speakers to use the expression in that particular way.

On their view, externalist approaches to metasemantics such as reference magnetism are ultimately unsatisfactory, for they render metasemantics explanatorily vacuous. If the speakers of a certain community treat a certain term *e* as if it were a natural kind term with referent *f* but the externally fixed metasemantics for *e* says that it is a functional kind term with referent *g*, then what we need to take into account in order to assess how people actually behave (i.e., their actual dispositions to correct their usage, how they manage to coordinate with each other, etc.) is still *f*, not *g* (see also Cohnitz & Harukioja 2013). The chapter closes by mentioning some consequences that this internalist metasemantic view has for empirical research. The authors argue, for example, that the conclusion that natural kind terms are used referentially on some occasions and descriptively on others, as some authors have argued on the basis of empirical data (cf. Nichols et al 2015), is too hasty.

Experimenters should test not just the initial responses of speakers, but also their dispositions to correct their usage, i.e., to acknowledge that they have made a mistake. Thus, this last chapter serves as a good conclusion for the first part of the book, which as we have just seen puts forward an important amount of challenges that any cautious researcher should bear in mind when carrying out empirical research in fields like semantics or pragmatics. Let us now review the second part.

Joshua Glasgow's chapter – the eight so far – investigates further into the relation between meaning and the dispositions of the linguistic community. And his main assumption is that meaning is revealed, precisely, by the dispositions of speakers to stop a dispute on the grounds that the disagreement is merely verbal. Although his view is broadly descriptivist – referents are fixed by a set of properties that something must uniquely have –, he claims that the classical objections to descriptivism do not apply to his theory, for what secures a common meaning is not sharing the same descriptions, but rather having the same set of dispositions to negotiate. He also mentions a central consequence of his theory for the topic of conceptual revolution: namely, that it occurs whenever the dispositions of the speakers change. Thus, he claims, any intentionally induced meaning shift (i.e., any project of conceptual engineering) should proceed by modifying the dispositions of the linguistic community.

The next three chapters focus on the folk concept of *race* and its relationship with scientific variants of the term as used in biology or social sciences. The contribution by Edouard Machery and Luc Faucher (chapter 9) presents the results of four empirical studies on whether race is ordinarily perceived as a social or as a biological category. And the outcome of the studies points towards the latter conclusion: people tend to induce biological and even psychological

traits of individuals on the basis of race much more frequently than they would infer them on the basis of prototypical social categories. This is in line with the idea that biological categories bear a rich inductive power (Gelman & Markman 1986). Although the experiments correctly test the inductive power of the folk concept of *race* in contrast with that of other folk concepts such as *gender* and *social class*, I believe that the studies would have benefited from comparing "race" with other, less loaded terms with a biological component and that may also suggest common ancestry, such as "hair color".

The tenth chapter, by Esa Díaz-León, examines the mismatch between the ordinary concept of *race* and its alleged referent. She builds on Glasgow's (2003) argument to the effect that, when searching for a scientifically backed referent for folk concepts like *race*, the meaning revisions that we are allowed to perform on the term cannot deviate much from its original sense, or else we would not be talking of race at all – if nothing matches the descriptions ordinary speakers associate with the term, then it must be treated as a case of failed reference, pretty much like "witch". Díaz-León attempts to provide an answer to the problem of how to decide between a theory that claims that a term like "race" is a case of failed reference and another one in which it does have a referent that, for some reason (say, because of reference magnetism), fails to satisfy most of the properties we usually associate with it. And her response is that we should let normative considerations guide us in these decisions. What counts as natural is not just there to be found, for there are many equally valid ways to carve the world at its joints, i.e., many possible divisions that may be better or worse suited for our projects. Our choice should be guided by the value of the projects they serve – which, of course, also includes the *ethical* value of those projects.

Chapter 11, by Robin Andreasen, argues that "race" is a fragmented concept, i.e., a term that started out having a univocal sense but ended up having several, diverging uses due to the variety of theoretical roles it is supposed to fulfill. More specifically, Andreasen argues that it is currently fragmented in at least two theoretical race concepts: a social category of *race* with no biological significance, and the notion of *race* as employed in racial population naturalism. In view of this duality, we have three options at hand: we may reject the use of a non-univocal term like "race", we may choose to selectively eliminate "race" in one sense but not in the other, or we may preserve both uses, each restricted to its own domain. And she rejects selective eliminativism on the basis that, if a term like "race" ought to be eliminated in one scientific area because of ethical considerations, then it should be eliminated from both. The concept of race as a social category is usually seen as a liberatory notion, but that does not mean that the kind of research in which it is employed – which makes generalizations on issues like health disparities and outcome inequalities – could not be misappropriated or distorted in favor of racist ideas. She also warns against the outright elimination of "race", for it is not clear whether the candidates to substitute it in scientific contexts ("ancestry", for example) are sufficiently devoid of negative connotations. In view of these issues, it appears that retaining a non-univocal concept of *race* does not seem so unreasonable.

In chapter 12, Sally Haslanger responds to several challenges that conceptual amelioration may face. On her view, individuating a concept requires not just focusing on its content, but also on the way in which it presents how the world is divided. Thus, two concepts like *water* and $H_2O$ share a single content, yet they differ in where they situate that very same extension: the former locates it within the space of beverages, whereas the latter does so within the domain of

chemicals. This means that these two concepts carve the world differently, and modifying one of them may be advisable even when modification of the other is not. Thus, epistemic amelioration occurs when new, better ways to access the content of a concept are disclosed, whereas semantic amelioration, in turn, can occur when the content shifts. Haslanger argues that none of these changes entail an actual shift of concept, rather than a mere modification of the same concept. To this effect she calls attention into the concept of *meter*, which has been modified in at least six different ways since it was first introduced, both semantically and metasemantically, in order to make it more precise. There is no reason why we should not be able to do the same with terms like "marriage".

Haslanger also addresses an argument by Hermann Cappelen (2018) to the effect that semantic amelioration is impossible, for the metasemantics is inscrutable and the process of change is incontrollable. She agrees with the latter idea, but encourages philosophers working on these issues to play a role in social movements by providing people new possibilities of cognitive access to those contents. Taken together, the chapters by Haslanger and Díaz-León point towards a joint conclusion: the task of the conceptual engineer is to provide new ways to carve the world, and to do so by considering which of them are more useful for socially valuable movements.

In the last chapter of the book, Teresa Marques raises a new issue for conceptual engineers to bear in mind: the very legitimacy of their ameliorative projects. A quick look on the way in which words like "honor" or "free elections" are used in totalitarian regimes is enough for showing that meaning revisions often result in *perversion*, rather than amelioration, and this means that, in order to tell whether a revisionary project is permissible, we need to make sure first that our project is not actually perverting the meaning of the

engineered concept. In the framework she employs (a Stalnaker-style (1978) two-dimensional semantics as presented in García-Carpintero & Marques (2019)), terms are endowed not just with their usual content, but also with the ability to trigger expressive presuppositions – that is how derogatory language works. And meaning perversions exploit precisely that: by calling a certain practice "democratic", the presupposition that it is positive is triggered, and it obliges those who confront such practice to justify that they are not opposing actual democracy.

As a corollary, Marques takes a skeptical stance towards the idea that we can ameliorate the meaning of terms like "race" just by shifting their extension. Whatever extension we may choose for the revised term, it is unlikely that the expressive presuppositions it bears will go away. The latter is a clear point of convergence with the ideas presented by several authors of this book. Compare with Andreasen: the expressive presuppositions that a term like "race" is endowed with are not likely to disappear just because the term receives a social category as its extension. And, pretty much like Haslanger, the ideas presented by Marques show that conceptual engineers should not operate just on the extension of the terms, but also on other aspects of their meaning.

All in all, I believe that the volume makes an important contribution to the interdisciplinary study of concepts. The first part of the book makes both negative and positive contributions: on the one hand, it raises important challenges that need to be faced when dealing with concepts, such as the vast number of ways in which they may display variation or the limitations of traditional philosophical approaches; on the other, it puts forward an important number of methodological *provisos* and suggestions that contribute to the construction of a new framework from which these issues can be addressed. The second part, in turn, establishes

very fruitful connections between several related debates: the relation between folk concepts and their scientific counterparts, the topic of diachronic variation in concepts, and the right procedures – both ethical and methodological – that should be followed when pursuing an intentionally induced meaning change. In this regard, it is worth mentioning that the several chapters of the book are very well-connected, specially in the second part, in which most authors cite each other's work. The introduction by the two editors also helps establish several links between the several contributed chapters. In sum, this is a very comprehensive collective work that addresses several connected topics from a wide range of perspectives and sets important grounds for future research.

## References

Cohnitz, D., & Haukioja, J. (2013). Meta-externalism vs meta-internalism in the study of reference. *Australasian Journal of Philosophy*, 91(3), 475-500.

Frege, G. (1892). Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik*, 100, 25–50.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23, 183–209.

Glasgow, J. (2003). On the new biology of race. *Journal of Philosophy*, 100(9), 456–74.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics, 3: Speech Acts* (pp. 41–58). New York: Academic Press.

Lewis, D. (1984). Putnam's paradox. *Australasian journal of philosophy*, 62(3), 221-236.

Nichols, S., Pinillos, A., & Mallon, R. (2015). Ambiguous reference. *Mind*, 125, 145–75.

Marques, T., & García-Carpintero, M. (2020). Really expressive presuppositions and how to block them. *Grazer Philosophische Studien*, 97(1), 138-158.

Marques, T., & Wikforss, Å. (Eds.). (2020). *Shifting Concepts: The Philosophy and Psychology of Conceptual Variability*. Oxford: Oxford University Press.

Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics*. New York: Academic Press.