

How words matter: A psycholinguistic argument for meaning revision

Steffen Koch 

Philosophy Department, Bielefeld University, Bielefeld, Germany

Correspondence

Steffen Koch, Abteilung Philosophie, Universität Bielefeld, Postfach 10 01 3, D-33501 Bielefeld, Germany.
Email: steffen.koch@uni-bielefeld.de

Funding information

Open Access funding enabled and organized by Projekt DEAL.

Linguistic interventions aim to change our linguistic practices. A commonly discussed type of linguistic intervention is meaning revision, which seeks to associate existing words with new or revised meanings. But why does retaining old words matter so much? Why not instead introduce new words to express the newly defined meanings? Drawing on relevant psycholinguistic research, this paper develops an empirically motivated, general, and practically useful pro tanto reason to retain rather than replace the original word during the process of conceptual improvement.

KEYWORDS

conceptual engineering, kinds, lexical effects, linguistic interventions, meaning revision, mutual exclusivity

1 | LINGUISTIC INTERVENTIONS AND THE PROBLEM OF MEANING REVISION

In philosophy and elsewhere, people often argue about language: about how we should refer to categories of race and gender, or sexual harassment, or COVID-19 restrictive measures, or tax cuts, or global warming. Sometimes, such disputes concern the linguistic status quo—whether, according to the *actual* use and meaning of “sexual harassment,” a given behavior qualifies as such. At other times, however, such disputes involve attempts to *change* our linguistic practices, such as changing the way we use gender terms like “woman,” stopping talking about “fake news” (Habgood-Coote, 2019), or replacing “social distancing” with “physical distancing” (Schnell, 2020). Borrowing a term from Sterken (2020), I will refer to such attempts as *linguistic interventions*.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Mind & Language* published by John Wiley & Sons Ltd.

This paper is concerned with linguistic interventions that target words and their meanings (as opposed to, e.g., sentences or non-meaning related aspects of language). Three basic types of such linguistic interventions can be distinguished:

Introduction: Attempting to introduce a new lexical item along with a new meaning m^* .

Revision: Attempting to change the meaning of an existing lexical item from m to m^* .

Elimination: Attempting to eliminate an existing lexical item along with its meaning m .

Introduction and elimination can be combined to yield:

Replacement: Attempting to replace an existing lexical item along with its meaning m by a different lexical item with a similar, though modified meaning m^* .

In these definitions, as in the rest of this paper, I take *lexical items* or *words* to be lexical forms that are not individuated by their associated meanings. In this sense, a given lexical item could possibly be paired with all sorts of different meanings. Moreover, following Fischer (2020), I will understand the *meaning* of a word as that which guides how speakers “assess the correctness of categorization judgments, the accuracy of descriptions using the word, and the validity of inferences from utterances using the word” (p. 4). As others have argued, it is this sense of meaning that is particularly relevant in the context of linguistic interventions (cf. Nado, 2023; Thomasson, 2021).

Revision is particularly prominent in the recent philosophical discussions of conceptual engineering (Cappelen & Plunkett, 2020; Isaac et al., 2022), where the idea is often that we should associate existing words with *new* or *revised* meanings.¹ But while conceptual engineering is currently very popular, it also faces challenges. Strawson (1963) objected early on that the re-engineering of concepts to solve philosophical problems threatens to change the subject rather than to provide a solution. Similarly, Cappelen (2018) observes that revision can easily lead to miscommunication and merely verbal disputes. Others argue that even well-supported conceptual engineering projects are very difficult to implement: Conventionalized meanings are recalcitrant (Koslow, 2022) and difficult to suppress (Fischer, 2020); depending on one’s meta-semantic background theory, meaning change may prove difficult to the point of infeasibility (Cappelen, 2018; Deutsch, 2020). According to Ritchie (2021a), when revision is applied to social categories, it often reinforces essentialist thinking rather than mitigating it.

In this paper, I argue that there can be valid reasons for retaining existing terminology through revision rather than introducing new terms through introduction, despite the difficulties involved. I am not the first to suggest this. Haslanger (2000) suggests that reappropriating existing terms for social groups through revision can shift expectations and inferences in ways that introduction cannot. Cappelen (2018, 2020) argues that words have lexical effects beyond their semantic and pragmatic effects, and that preserving these effects may provide a reason to prefer revision over introduction and replacement (see also Landes, forthcoming). Others argue that certain cases of meaning change can preserve topic continuity (e.g., Belleri, 2021; Cappelen, 2018; Flocke, 2021; Knoll, 2021), and that preserving lexical continuity can help

¹Some recent examples: Haslanger (2000) suggests changing our use of “woman” to refer to (roughly) systematically oppressed female persons; Dembroff (2016) suggests using “sexual orientation” to refer to a person’s sexual preferences without mention of their own sex or gender; Manne (2017) suggests using “misogyny” to refer to hostile social forces faced by girls and women that serve to enforce a patriarchal order; Fassio and McKenna (2015) suggest using “knowledge” to refer to an epistemic state of stake-sensitive invariantism.

demonstrate that the ameliorated meaning remains connected to ongoing conversations related to the original word (Cappelen, 2020, p. 144).

The primary goal of this paper is to provide a novel rationale for revision. The account I develop here stands out from existing proposals in that it does not rely on a controversial concept of topic continuity² and is based on solid empirical findings rather than a priori reasoning. Although I consider these features to be improvements over current accounts, my primary goal is not to critique these other accounts, but rather to introduce a new one.

The account I develop draws on two well-established psycholinguistic findings about word processing and word learning: the *noun-kind bias* and the *mutual exclusivity bias*. I argue that these biases not only explain why revision has different psychological effects on the target group than either introduction or replacement, but they also reveal why, under appropriate circumstances and given plausible goals, revision has clear advantages over introduction and replacement. Whether these advantages ultimately outweigh possible disadvantages of revision will vary from case to case. My goal, then, is not to provide a universally compelling *pro toto* reason for revision, but a *pro tanto* reason that can be counterbalanced by contrary considerations. This aspect of my account is not a disadvantage; rather, it recognizes that terminological choices are inherently complex and that a simple “one-size-fits-all” solution is unrealistic.

The account offers three important benefits. First, its conclusion is based on robust *empirical evidence* about word learning and processing. Second, the account provides a set of *general criteria* for determining when revision is beneficial. It encompasses a wide range of thematically diverse cases, making it applicable to numerous actual and potential linguistic interventions. Third, since it is typically possible to assess whether a given case meets the criteria outlined in this account, it provides *practical guidance* for would-be linguistic interveners. In sum, the account presents a novel, empirically grounded, and generalized rationale for retaining existing terminology that promises to be valuable in practice.

The remainder of this paper is organized as follows. Section 2 presents the psycholinguistic background of the account. It introduces the mutual exclusivity bias and the noun-kind bias and outlines some of the evidence for the existence of each of them. Section 3 draws implications of these biases for would-be linguistic interveners. Using Clark and Chalmers's discussion of “belief” as a case study, it shows why these two biases often support revision. Section 4 generalizes the lesson, while Section 5 discusses its limitations.

2 | PSYCHOLINGUISTIC BACKGROUND

2.1 | Nouns and kinds

The first psycholinguistic finding that I will use in my case for revision concerns the relationship between kind representations and noun use. Psychologists have amassed a substantial body of evidence suggesting that people distinguish between categories they consider to be *kinds* and other types of categories. Categories perceived as kinds are characterized by their *cohesiveness*, *inductive potential*, *informativeness*, *relatively stable membership*, and their role in certain forms of *explanation* (Gelman, 2003; Markman, 1989; Prasada et al., 2012).

Paradigmatic examples of categories represented as kinds include supposedly natural kinds such as *tigers*, *water*, *electrons*, and *elm trees*. Tigers, for example, are considered a

²For criticism, see Koch (2023) and Kocurek (2022).

cohesive category whose members share numerous properties. This is why, upon learning that one particular tiger has stripes, we infer that other tigers also have them. Similarly, discovering that something is a tiger provides us with a wealth of additional information about it. Furthermore, we consider being a tiger to be a stable property: We believe that tigers rarely stop being tigers, and that if you are not one already, there is little you can do to become one. Finally, we assume that something being a tiger explains many of its properties, including its behavior.

What applies to natural kinds like tigers also applies to social kinds like *women*, *philosophers*, or *scientists* (Diesendruck et al., 2013; Rhodes & Gelman, 2009), and even to artifact kinds like *cups* or *chairs* (Keleman & Carey, 2007). People tend to think of *being a woman* as more cohesive, inductively potent, informative, stable, and explanatory than *having a weakness for licorice*. Of course, most human categories are seen as less stable than being a tiger. Few people believe that one is born a philosopher, and gender is considered to be subject to change. Nevertheless, compared to other properties, such as *sitting at a desk right now* or *wearing a gray shirt*, these properties seem relatively stable and seem to provide substantial insight into the individual in question.

Crucially for the argument to follow, there is strong evidence for a connection between being perceived as a kind and the use of nouns. Categories denoted by nouns are more likely to be perceived as kinds and, consequently, as being inductively potent, cohesive, informative, and so forth. (Foster-Hanson & Rhodes, 2020; Gelman & Markman, 1986; Markman, 1989).³ This applies in contrast to both longer phrases and adjectives. When a category is identified by a lexical nominal rather than a phrasal nominal, people are more likely to describe and understand it as a “single kind of thing” (Prasada et al., 2012). Also, when the same noun is used to describe two things, children are more likely to judge them as similar than when the same adjective is used (Gelman & Coley, 1990). Furthermore, studies indicate that even adults perceive preferences and traits as more persistent and stable when described with nouns rather than adjectives (Walton & Banaji, 2004).

In a well-known study, Gelman and Heyman examined how 5-year-olds and 7-year-olds judge the stability of a property, its past and future behavior, and its counterfactual behavior, depending on whether it is described by a predicate nominal or a predicate adjective. The study revealed that children perceived properties as more persistent and stable when predicate nominals were used. Because the nouns in the study had not been previously lexicalized (e.g., “carrot-eater”), the authors concluded that “children were not retrieving rote meanings, but rather made use of a general rule that they applied to these novel noun phrases” (Gelman & Heyman, 1999, p. 491). Ritchie (2021b) replicated these findings with adults using entirely new and invented expressions such as “dax.”

In sum, there is a significant psychological difference between representations of mere categories and those of genuine kinds. In general, when a category is represented as a kind, it is thought to have greater inductive potential, to be more informative, to have relatively stable membership, and to play a role in certain forms of explanation. Kind representations are not exclusive to natural world categories; they can also apply to social and artifact categories. Furthermore, the use of nouns rather than adjectives or longer phrases strengthens the tendency to represent a given category as a kind.

³For language effects pertaining to the use of generic expressions such as “Mosquitos carry the West Nile virus,” see, for example, Leslie (2017) and Wodak et al. (2015).

2.2 | Mutual exclusivity

How do children and adults learn new word meanings? This question has preoccupied psychologists almost from the beginning of the discipline. It received increased attention from both psychologists and philosophers in the 1960s, when Quine used his “gavagai” example to illustrate the problem of referential ambiguity (Quine, 1960). One of the most widely discussed phenomena in the word-learning literature is *mutual exclusivity*: When confronted with a previously unfamiliar word, language learners seem to assume that the meaning of that word refers to a novel category—a category for which they do not yet have a name. This bias helps children and other language learners to integrate unfamiliar expressions into their stock of known word meanings.

Evidence for mutual exclusivity dates back to the early 1980s. In an original study by Kagan (1981), children were shown two familiar objects (a doll and a dog) and an unfamiliar object. After playing with them for a while, they were presented with an unfamiliar word, as in “Give me the zoob.” Children as young as 22 months tended to choose the unfamiliar object. This suggests that they believed that a novel word did not refer to objects that already had names. This finding was replicated in later studies by Markman and Wachtel (1988), who also found that when only familiar objects are present, children tend to think of the novel word as referring only to a part of that object rather than to the object itself. Since then, these findings have been repeatedly replicated and extended along various dimensions (e.g., Au & Glusman, 1990; Grassman et al., 2015; Hall, 1991; Hutchinson, 1986; Merriman & Bowman, 1989).

The most striking extensions are, first, that mutual exclusivity holds not only for common names, but also for proper names (Diesendruck, 2005), adjectives (Carey & Bartlett, 1978) and verbs (Merriman et al., 1996); and second, that it holds not only hold for children, but for anyone trying to figure out the meaning of an unfamiliar word. Golinkoff et al. (1992) tested mutual exclusivity with adult university students. After presenting them with three objects for which they already had a name and one for which they did not (e.g., a broom handle tip for hanging a broom on a nail, or rubber feet for the bottom of chairs), they used a newly invented expression (e.g., “dax,” “jick,” and “zorch”) to test which object the students would assign it to. 100% of the participants assigned the unfamiliar word to the previously unnamed object (see also Halberda, 2006). The effect sizes of these experiments led Bloom (2000) to conclude: “Mutual exclusivity is not a subtle phenomenon: you don’t need to test dozens of children in careful laboratory conditions to see it at work” (p. 67).

It is not really a question, then, whether mutual exclusivity is a real phenomenon. Mutual exclusivity is one of the key factors in limiting an otherwise insurmountable number of possible meanings for each unfamiliar word that a language learner is trying to learn. It is also consistent with, and indeed explains, other peculiarities of language development, such as children’s tendency to stop calling objects by incorrect names once they have learned the correct ones (even without explicit correction) (Clark, 1987), or children’s difficulties in learning superordinates (Macnamara, 1982). Children, and language learners in general, are reluctant to accept more than one label for each object (or category). However, mutual exclusivity is a bias or “default assumption” (Liittschwager & Markman, 1994), not an immutable constraint. Words with overlapping references *can* be learned—they are just *harder* to learn (more on this in Section 4 below).

The nature and origin of the mutual exclusivity bias is more controversial among psychologists than its existence. Until the early 2000s, three main hypotheses prevailed. Mervis et al. (1994) proposed that mutual exclusivity is a specifically lexical phenomenon, a feature of how words work that is either innate or acquired during early language development. Markman (1992), on the other hand, suggested that mutual exclusivity is a special case of a more general

learning principle that leads children to prefer one-to-one mappings as part of a general tendency to expect regularities. Finally, Bloom (2000) argued that mutual exclusivity results from people's theory of mind, which serves as a means of interpreting the likely speaker's meaning of utterances involving unfamiliar expressions. Recent trends in the word-learning literature suggest that mutual exclusivity results from a combination of various domain-general processes, such as attentional biases working together (Kucker et al., 2018; Samuelson & McMurray, 2017). While this discussion is both important and intriguing, this paper does not take a position on the underlying nature of mutual exclusivity.

In summary, it is well established that language learners exhibit mutual exclusivity: a bias to treat word meanings as referring to non-overlapping categories. This bias is found at all ages and with respect to different lexical types, such as nouns or verbs. One consequence of this bias is that people find it significantly more difficult to learn words that refer to objects for which they already have words, and that in such cases they are more likely to misinterpret a given utterance.

3 | IMPLICATIONS FOR LINGUISTIC INTERVENTIONS: THE CASE OF “BELIEF”

I will now argue that the combination of the above biases—the bias to interpret nouns as referring to kinds and the mutual exclusivity bias—together provide a pro tanto reason for favoring revision over introduction and replacement, at least when certain conditions hold. The central idea of the account is that, when these conditions hold, people's linguistic biases are such that keeping the old word has better—more goal-conducive—cognitive effects on the target group than switching to a different word. To make the discussion more concrete, I will develop the argument using the example of “belief”; later on, I will discuss how the lesson generalizes to other cases and discuss an important limitation.

According to the famous extended mind hypothesis of Clark and Chalmers (1998), mental properties and states are not necessarily limited to what happens inside the skull, but can be extended to external devices such as notebooks or smartphones. One implication of this view is that a person can believe a piece of information even if it is not stored in her biological memory, as long as she can easily and reliably retrieve it. For example, I can be said to believe that the Globe Theatre in London is located near Southwark Bridge station not because I am currently aware of this fact or have it stored in my biological brain, but simply because I carry a smartphone that will spit out this information as soon as I ask for it.

Following Cappelen (2018, Chapter 1), I will treat Clark and Chalmers's proposal as an instance of conceptual engineering; more specifically, as an attempt to revise the meaning of “belief” to include externally stored, yet easily and reliably accessible information. Clark and Chalmers' justify their proposal as follows:

We do not intend to debate what is standard usage; our broader point is that the notion of belief *ought* to be used so that Otto qualifies as having the belief in question. In all *important* respects, Otto's case is similar to a standard case of (non-occurrent) belief. The differences between Otto's case and Inga's are striking, but they are superficial. By using the “belief” notion in a wider way, it picks out something more akin to a natural kind. The notion becomes deeper and more unified, and is more useful in explanation.

(Clark & Chalmers, 1998, p. 14)

At this point it will be helpful to introduce some terminology. Let us use the label *inliefs* to refer to all those beliefs that are stored in one's biological memory, that is, *inside* the skull; let us also call beliefs that are stored on *external* devices but are still easily and reliably accessible *exliefs*; finally, let us call the set union of *inliefs* and *exliefs*, that is, *all* candidate beliefs together, *alliefs*.⁴

Couched in this terminology, Clark and Chalmers argue that, regardless of whether “beliefs” currently refers to *alliefs* or just *inliefs*, it *should* refer to *alliefs* because the differences between *inliefs* and *exliefs* are “superficial” and because *alliefs* are “more akin to a natural kind” that is “more useful in explanation.” So even if *exliefs* are *not* covered by the standard usage of “belief,” they *should* be. In the discussion that follows, I will simply assume that *exliefs* are *not* covered by the standard meaning of “belief” and that their proposal thus involves an attempt to *change* the standard meaning of “belief.”⁵

Clark and Chalmers's proposal consists of two claims, one that is explicitly defended and one that is implicit. Their explicitly defended claim is that *alliefs* constitute a more useful and more natural category than the narrower category of *inliefs*, so we should use the concept of *alliefs* when thinking and theorizing about the mind. The more implicit claim is that we should use the word “beliefs” and its cognates to refer to the category of *alliefs*. In what follows, I will simply assume that Clark and Chalmers' first claim is well defended, and focus instead on the second. The question I will address, then, is: Given that the newly identified concept of *alliefs* is superior to the concept of *inliefs*, why and under what circumstances should we use the word “beliefs” to refer to this new concept? My strategy for answering this question will be to lay out the different terminological options and to compare their putative cognitive effects. Here, the two biases introduced in the last section will be put to work.

The peculiarities of Clark and Chalmers' view impose constraints on their choice of terminology. In particular, an appropriate choice of labels should be such that:

- (i) It encourages people to treat *alliefs* as a natural kind.
- (ii) It does not encourage people to treat *inliefs* as a natural kind.
- (iii) It does not encourage people to treat *exliefs* as a natural kind.
- (iv) It does not encourage people to think that *inliefs* and *alliefs* are mutually exclusive categories.

Where do these constraints come from? Clark and Chalmers write that *alliefs* are “more akin to a natural kind” and that the notion is “deeper,” “more unified,” and “more useful in explanation” than *inliefs* and presumably *exliefs*, too. Here I make the simplifying assumption that, according to Clark and Chalmers, *alliefs* are a natural kind, whereas *inliefs* and *exliefs* are not. To be sure, the phrase “more akin to a natural kind” does not imply such a commitment and as I will show later on, the comparative assessment implied by this phrase is ultimately sufficient for the argument (see Section 4). Now if, as the simplifying assumption suggests, Clark and Chalmers are indeed of the view that *alliefs* are a natural kind and that *inliefs* and *exliefs* are not, then surely they would want to choose terminology that encourages people to treat these categories in just this way. For this reason, (i) to (iii) seem reasonable. Moreover, according to the definition proposed by Clark and Chalmers, *inliefs* are a subset of *alliefs*, so they are clearly

⁴What I call “alliefs” is not to be confused with what Tamar Gendler calls “Aliefs” (Szabó Gendler, 2008).

⁵Chalmers (2020) expresses sympathy with this interpretation, but notes that, according to their own view back in 1998, “extended cases of beliefs were literally beliefs” (p. 8).

not mutually exclusive categories. Since there is no reason to encourage people to think otherwise, (iv) also seems reasonable.

With these constraints in play, we can now look at the terminological options available for framing the extended mind hypothesis. There are, of course, an infinite number of such options, but let us consider only those that provide at least adequate *expressive power*, that is, the linguistic tools to refer to the things worth referring to (as far as Clark and Chalmers are concerned). This reduces the number of options to the following four:

- (a) *Introduction 1*: Introduce a new word for *exliefs* (e.g., “exliefs”) and retain “belief” with its current meaning.
- (b) *Introduction 2*: Introduce a new word for *alliefs* (e.g., “alliefs”) and retain “belief” with its current meaning.
- (c) *Replacement*: Introduce a new word for *alliefs* and eliminate “belief.”
- (d) *Revision*: Revise “belief” to refer to *alliefs* instead of *inliefs*.

The option favored by Clark and Chalmers is (d). I will now argue that, given the constraints (i) to (iv) above, and judged from the perspective of mutual exclusivity and the noun-kind bias, (d) is indeed the best choice, since it has considerable advantages over the other candidates.

Let us begin with (a). What happens if we simply introduce a new word, for example, “exliefs,” to refer to the *hitherto* unnamed category of *exliefs* and retain “belief” with its current meaning? Since, by the assumptions made above, our current word “belief” refers only to *inliefs*, this enables us to refer to both *inliefs* and *exliefs*, albeit with separate words. The requirement of adequate expressive power is therefore satisfied. However, the discussion of the noun-kind bias in the last section showed that people are likely to treat nouns as referring to (natural) kinds. Therefore, option (a) will promote thinking of *inliefs* and *exliefs* as constituting (natural) kinds. But this violates conditions (ii) and (iii) above, according to which people should not be encouraged to treat either *inliefs* or *exliefs* as natural kinds. This aspect makes option (a) a terminological choice that is seriously misleading. Moreover, since option (a) does not reserve a noun for the set union of the extensions of *inliefs* and *exliefs*—the broader category of *alliefs*—it does not encourage people to treat this category as a natural kind, which stands in tension with constraint (i) above. In the absence of strong conflicting considerations, option (a) does not seem to be a promising candidate.

What if instead we introduce a new word to refer to the larger category of *alliefs* (e.g., the word “alliefs”) and retain “belief” with its current meaning, that is, option (b)? The result is a situation in which we are able to refer to *alliefs* (via “alliefs”) and to *inliefs* (via “beliefs”). Again, the requirement of adequate expressive power is satisfied. Moreover, what is good about this strategy is that, due to the noun-kind bias, it encourages people to treat *alliefs* as a (natural) kind, thereby satisfying constraint (i) above. This is an advantage that (b) has over (a). Nevertheless, (b) has two significant disadvantages. Since the word for *inliefs* (“belief”) has not been eliminated, the noun-kind bias tells us that option (b) encourages people to treat *inliefs* as a natural kind, which violates constraint (ii). Furthermore, the mutual exclusivity bias tells us that, as long as there is a word for *inliefs* (“beliefs”) and an additional word for *alliefs* (“alliefs”), people will tend to think that the category of *alliefs* does not overlap with the narrower category of *inliefs*, which is false and violates constraint (iv) above. In light of this (and again in the absence of conflicting considerations), option (b) does not seem to be a good strategy either.

Can we improve on option (b) by eliminating the word for *inliefs* (“beliefs”)? This is what option (c) suggests. It results in a situation where there is only one word to refer to believe-like

states: the word “alliefs,” which refers to *alliefs*. This option is a significant improvement over (a) and (b), since it does not encourage treating *inliefs* and *exliefs* as natural kinds (and is thus consistent with constraints (ii) and (iii)). Moreover, since it reserves a noun for *alliefs*, it encourages treating *alliefs* as a natural kind (and is thus consistent with constraint (i)). But there is one remaining problem with this option. By the time that the new word “alliefs” is introduced and spreads through the linguistic community, people will still be familiar with “beliefs.” For this reason, and in the absence of any explicit corrections, they are likely to be guided by mutual exclusivity and thus interpret “alliefs” as referring to a category that does not overlap with the category referred to by “beliefs.” Like option (b), therefore, option (c) conflicts with constraint (iv).⁶ In the absence of overriding conflicting considerations, this makes option (c) a less than ideal choice.

Finally, let us take a closer look at option (d)—expanding the meaning of “beliefs” so that it refers to *alliefs* instead of just *inliefs*. First, this option satisfies the requirement of adequate expressive power, since it allows us to talk about all things worth talking about. Moreover, it is consistent with constraints (i), (ii), and (iii): It does not encourage people to treat *inliefs* or *exliefs* as (natural) kinds, because it does not reserve nouns for them; but it does encourage people to treat *alliefs* as a natural kind, because it reserves a single noun for this category. So up to this point, (d) has all the merits of (c). But unlike (c), (d) does not encourage false assumptions about the relationship between *inliefs* and *alliefs*. This is because it does not require the introduction of new vocabulary. Thus, holding the plausible constraints (i)—(iv) above in place, and judging only by the likely cognitive effects of the noun-kind bias and the mutual exclusivity bias, (d) appears to be the best option of the lot. From this perspective, Clark and Chalmers’s second thesis is defensible after all: It does indeed make sense to revise the use (or meaning) of “belief” to refer to the larger category of *alliefs*.⁷

4 | GENERALIZING THE LESSON: FROM “BELIEF” TO MEANING REVISION

What applies to the case of “belief” also applies to many other cases. This subsection serves to generalize the lesson by showing that the argument given in the above applies to all cases with a certain general structure. This generalization will also show that Clark and Chalmers’s comparative claim that the preferred category is “more akin” to a natural kind rather than definitely constituting one, is sufficient to get the argument going.

Clark and Chalmers’s take on “belief” instantiates a *category-expansion project*: a proposal to exchange a given category for a larger, more inclusive category. It is easy to see that the argument given in the above applies to many other category-expansion projects as well. In particular, the argument applies to any proposal to swap some category c_1 for some other category c_2 , provided that the following four conditions apply:

- (1) c_1 is named by some word w , whereas c_2 is unnamed.
- (2) c_1 is a subset of c_2 .

⁶There is an important exception to this claim that will be discussed in Section 5 below.

⁷It is consistent with this proposal that one uses adjectives instead of nouns to distinguish between the different subcategories of beliefs, as in “internal beliefs” and “external beliefs.” Since adjectives are not within the scope of the noun-kind bias, this does not fall prey to the criticism raised against options (a) to (c) above.

- (3) We should treat c_2 as a (natural) kind.
- (4) We should not treat c_1 as a (natural) kind.

These conditions were satisfied by the “belief” case, but they are satisfied by many other cases as well. Consider, for example, what has happened (or is happening) to the meanings of words like “family,” “intelligence,” “work,” “health,” or “education.” All of these words seem to have had rather specific meanings, but have now broadened their applications, for example, from just *physical health* to include *mental health*. These cases stand a good chance of satisfying the above conditions. Whenever they are satisfied, the noun-kind bias makes it unwise to keep a noun for c_1 but supports using one for c_2 . And since c_1 is a subset of c_2 , the mutual exclusivity bias makes it *prima facie* misleading to introduce a new word for c_2 .

Interestingly, however, we can generalize even further, beyond category-expansion projects bound by condition (2). To see this, consider *category-refinement projects* where the order of c_1 and c_2 is reversed, that is, where the new category c_2 is a subset of c_1 . The IAU’s redefinition of “planet” is an example: Whereas the term used to include all celestial bodies orbiting the Sun, its revised meaning requires that these bodies also be nearly spherical in shape and have their orbits cleared of debris. As long as such projects satisfy (1), (3), and (4) above, the argument applies to them as well. Again, the noun-kind bias makes it unwise to keep a noun for c_1 but supports using one for c_2 . And since c_2 is a subset of c_1 , the mutual exclusivity bias again makes it *prima facie* misleading to introduce a new word for c_2 . So, we can generalize (2) to (2*):

(2*) c_1 is a subset of c_2 or vice versa.

In fact, the lesson generalizes even to cases where the extension of neither category is a subset of the respective other’s. As long as there is a significant degree of overlap between them, the argument still holds. Call projects that exhibit this general structure *category-shifting projects*. For an example of a category-shifting project that is neither expansion nor refinement, consider the shift in the meaning of “parent,” from referring to direct *genetical progenitors* to *primary caregivers* (cf. Haslanger, 2012, Chapter 5). Whenever we engage in a category-shifting project that satisfies conditions (1), (3), and (4) above, the noun-kind bias speaks against retaining a noun for c_1 and supports using one for c_2 , and the mutual exclusivity bias tells us that it is *prima facie* misleading to introduce a new word for c_2 . So, we can generalize (2*) to (2**):

(2**) There is considerable overlap between c_1 and c_2 .

How much of an overlap between c_1 and c_2 is necessary to satisfy (2**)? It is difficult to give a general answer to this question. If the overlap is too small or insignificant, the mutual exclusivity bias will work in the opposite direction: People will be better off treating c_1 and c_2 as two distinct and non-overlapping categories. Sufficient overlap is partly a matter of *quantity*: Whatever terminological choice we make should be the best one for as many applications of the word as possible. But it is also arguably a matter of *priority*: The terminological choice we make should be the best one for those applications that we prioritize. Trade-offs between quantity and priority are to be expected. In some cases, triggering correct assumptions about a few selected cases of high priority justifies the risk of triggering incorrect assumptions about many others. This will depend on the details of a given case and cannot be determined from the armchair.

Let us now turn to (3) and (4). Attributing instances of these claims to Clark and Chalmers in the last section proved to be difficult, since the authors never explicitly write that the

category they advocate, *alliefs*, is definitely a natural kind. However, we can now see that the argument for revision over introduction and replacement does not rest on these claims, for we can generalize (3) and (4) to just (3*):

(3*) We should treat c_2 as being closer to a (natural) kind than c_1 .

If (3*) is satisfied, our terminological choices are better if they encourage people not to treat c_1 and c_2 as being on a par with respect to kindhood. One way to satisfy this constraint is to choose words that encourage people to treat c_2 as a (natural) kind, but that do not encourage them to treat c_1 as such, for example, by using a simple noun to refer to c_2 but not to c_1 . Importantly, this holds even if c_2 ultimately fails to constitute a kind. All that really matters is whether people are better off treating c_2 as a kind than either just c_1 , or both c_1 and c_2 . For this to be the case, it is enough that c_2 comes closer to constituting a kind than c_1 . If this is true, then the argument covers to Clark and Chalmers's proposal, even though they stop short of claiming that *alliefs* definitely constitute a natural kind.

Finally, let us consider whether (3*) implies any weighty metaphysical commitments about kindhood. According to Bird (2018), *kind realism* is the view that “there are natural divisions among things such that our actual categorizations can succeed [or fail] in matching those divisions” (p. 2).⁸ Does (3) or (3*) imply kind realism? No. The above account imports its notion of kind from psychology. Psychologists use this notion to pick out categories that are thought to be particularly *cohesive*, to have inductive potential, to be informative, to have relatively stable membership, and to figure in certain forms of explanation. Thus, for kind representations to be accurate, it is not necessary that there be actual joints in nature or natural divisions of things. It is sufficient that categories may have the properties on the above list to varying degrees: that some categories are indeed more cohesive, have more inductive potential, are more informative, have more stable membership and figure in more or different types of explanations than other categories. All of this is compatible with the idea that how much inductive potential a category has or how informative it is ultimately depends on the contingent interest of human beings rather than the world itself.

To sum up, the pro tanto reason in favor of revision developed in the last section applies to any proposal to exchange some category c_1 for some other category c_2 that satisfies the following refined set of conditions:

(1**) c_1 is named by some word w , whereas c_2 is unnamed.

(2**) There is considerable overlap between c_1 and c_2 .

(3**) We should treat c_2 as being more cohesive, as having more inductive potential, as being more informative, as having more stable membership and as figuring in more types of explanations than c_1 .

With these generalizations in play, the account given provides a fairly general pro tanto reason to favor revision over introduction and replacement—a reason that we should expect to cover a great many of cases.

That said, it must be emphasized that conditions (1**) to (3**) should not be read as providing necessary and jointly sufficient conditions for when revision is the best option of the lot.

⁸Bird calls this view *weak realism* and contrasts it with *strong realism*—the view that “there are entities that are the natural kinds” (Bird, 2018, p. 8).

As mentioned in Section 1, terminological choices are complex and constrained by many different factors, some of which are case-specific. Moreover, the account presented here assumes that linguistic interveners will want to implement their respective proposals in a way that ensures smooth and successful communication. However, as Sterken (2020) argues, sometimes linguistic interventions aim instead to deliberately *disrupt* the interpretive common ground and the normal functioning of language in order to provoke metalinguistic reflection and thereby to effect changes in (linguistic) behavior (pp. 421–422). Depending on the case, this consideration may pull in either direction: to retain a given lexical item, even though this triggers false assumptions about its meaning, or to replace it, even though one is aware of the disruption that this may cause.

5 | OBJECTION: CAN MUTUAL EXCLUSIVITY BE SUSPENDED?

The noun-kind bias and the mutual exclusivity bias are biases, or tendencies, rather than necessities. The evidence discussed in the above shows that they are quite strong (the effect sizes are typically large, and they hold across populations and ages). Nevertheless, it is clear that both of them can be suspended or overridden. This is most obvious in the case of mutual exclusivity. If people were unable to suspend mutual exclusivity, they could hardly ever learn superordinate terms, and thus they would be unable to identify one single object as both a tiger *and* an animal, for example. This raises the question: If mutual exclusivity can easily be suspended, then why should we give it any weight in our terminological choices?

To answer this question, it is necessary to consider the conditions under which mutual exclusivity can typically be suspended. Existing research suggests that there are three such conditions: (i) there is evidence that one of the two words is a superordinate of the other (Au & Glusman, 1990); (ii) there is evidence that the two words belong to different languages (Au & Glusman, 1990); and (iii) there is explicit knowledge about how the meanings of the two words are related (Mervis, 1987). When it comes to linguistic interventions, (ii) clearly does not apply: The idea is to intervene in a single language, not to switch languages altogether. Except in special circumstances, (i) does not apply either. This leaves us with (iii), which we will take a closer look at.

Whether the target group can rely on explicit semantic knowledge of the kind required is likely to vary from case to case. In cases where the scope of the linguistic intervention is sufficiently broad and the target group rather heterogeneous, we should not expect large proportions of them to have explicit knowledge about the intended meaning and how it relates to existing meanings. The arguments and explanations from linguistic interveners have limited reach. It is unrealistic to expect them to reach a majority of language users within a sufficiently large target group. This means that if large-scale linguistic interventions are successful at all, it is usually not as a result of explicit teaching or persuasion, but rather as a result of changing conventions and social or moral norms that operate below the surface of our awareness.⁹ In such cases, it

⁹This point is convincingly argued by Thomasson (2021) and Nimtz (2021). According to Nimtz, manipulations of Bicchieri-style social norms (e.g., Bicchieri, 2006) constitute the most promising strategy of implementing conceptual engineering, precisely because it does not solely rely on convincing a majority of language speakers. I agree with Nimtz' assessment. The pro tanto reason for revision spelled out in this paper complements this approach, as it is especially pronounced in cases where the target audience cannot be reached via explicit arguments and definitions.

seems that—other things being equal—the most promising strategy is to make terminological choices that *exploit* rather than *counteract* the mutual exclusivity bias.

However, in cases where the scope of a linguistic intervention is much smaller and the target group more homogeneous, the situation may be different.¹⁰ For example, when an expert introduces a new technical term that is to be used only within a small expert community, the transfer of explicit knowledge about how its meaning relates to the existing vocabulary will usually work better. Although there are no controlled studies to prove this, it seems reasonable to assume that here the advantages of revision over replacement are *prima facie* less pronounced than for large-scale linguistic interventions. For example, if a linguistic intervention targets only academic philosophers with a certain specialization, it might be possible to provide them with the kind of information they need to suppress mutual exclusivity. Note that this does not yet imply that replacement is preferable to revision in such cases—this will largely depend on further case-specific considerations.

Finally, let me reiterate that the argument presented here is limited by the constraints developed in the last section. Many cases in which philosophers or other scientists introduce new technical vocabulary fall outside the scope of the argument. For example, technical terms such as “supervenience,” “grounding,” “rigid designator,” or “implicature” are not (and were never intended to be) category-shifting projects: projects that replace a given category with a closely related but more refined or useful one. The success of introducing such terms is not in tension with the argument presented here. As I have tried to show, recent projects concerning “belief,” and arguably also “family,” “health,” “woman,” or “planet” are relevantly different from those cases.

6 | CONCLUSION

The aim of this paper has been to propose a novel and empirically grounded *pro tanto* reason for preferring revision to other kinds of linguistic intervention. The starting point of my argument was two well-studied biases surrounding word learning and processing. According to the noun-kind bias, there is a tendency to interpret nouns as referring to genuine kinds, which are thought to be particularly cohesive and informative, to have inductive potential and relatively stable membership, and to figure in certain forms of explanation. According to the mutual exclusivity bias, people are likely to interpret unfamiliar words as having extensions that do not overlap with the extensions of familiar words.

I have argued that the existence of these biases has two important implications: First, that we should not introduce new nouns or keep old nouns in our language that do not refer to genuine kinds. Second, that we should not introduce new words whose extensions largely overlap with the extensions of words we already have at our disposal (unless, of course, there are good reasons for having both words).

Based on a case study that was later generalized, I have argued that these two upshots speak in favor of revision in cases that meet the following three conditions: (1**) the former category c_1 is named by some word w , while the proposed replacement category c_2 is as yet unnamed, (2**) there is considerable overlap between c_1 and c_2 , and (3**) we should treat c_2 as being more cohesive, as having more inductive potential, as being more informative, as having more stable

¹⁰Thanks to an anonymous reviewer of this journal for pressing me on this.

membership and as figuring in more types of explanations than c_1 (the notion of kindness that is salient in psychology).

I have also outlined some limitations of the account. Because the account focuses on facilitating a seamless transition between an old and a new or revised meaning, its conclusion is inapplicable if a linguistic intervener deliberately seeks disruption rather than smooth transition. Moreover, the mutual exclusivity bias has a less pronounced impact when language learners are given explicit definitions or explanations for a particular meaning. Although I have argued that such conditions are comparatively rare, we might find them, for example, in small expert communities. Finally, it cannot be emphasized enough that the arguments presented here detail one of several relevant considerations for arriving at well-informed terminological choices. These limitations aside, the account developed here provides a novel, empirically motivated and general pro tanto reason for the comparative value of revision over introduction and replacement that stands a good chance of being useful in practice.

ACKNOWLEDGEMENTS

This paper has long been in the making, and I am grateful to all the people who have helped along the way. In particular, I would like to thank Delia Belleri, Mirela Fus, Manuel Gustavo Isaac, Sigurd Jorem, Guido Löhr, Christian Nimtz, Jakob Ohlhorst, Mark Pinder, and Julia Zakkou for their feedback. I would also like to thank the audiences at the ESDIT 2022 Conference in Leiden, the GAP.11 in Berlin, the Conceptual Engineering: Normativity and Feasibility online conference in Bielefeld, the ICCE online conference, the members of the EXTRA.6 research colloquium in Bochum, and the members of the CONCEPT brownbag series in Cologne. Last but not least, thank you to two helpful and constructive reviewers of this journal. Open Access funding enabled and organized by Projekt DEAL.

DATA AVAILABILITY STATEMENT

There is no data available.

ORCID

Steffen Koch  <https://orcid.org/0000-0002-9360-799X>

REFERENCES

- Au, T. K., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: To honor or not to honor? *Child Development*, 61(5), 1474–1490.
- Belleri, D. (2021). Downplaying the change of subject objection to conceptual engineering. *Inquiry*, 1–24. <https://doi.org/10.1080/0020174X.2021.1908161>
- Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.
- Bird, A. (2018). The metaphysics of natural kinds. *Synthese*, 195(4), 1397–1426. <https://doi.org/10.1007/s11229-015-0833-y>
- Bloom, P. N. (2000). *How children learn the meaning of words*. MIT Press.
- Cappelen, H. (2018). *Fixing language. An essay on the foundations of conceptual engineering*. Oxford University Press.
- Cappelen, H. (2020). Conceptual engineering. The master argument. In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual engineering and conceptual ethics* (pp. 132–151). Oxford University Press.
- Cappelen, H., & Plunkett, D. (2020). A guided tour of conceptual engineering and conceptual ethics. In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual engineering and conceptual ethics* (pp. 1–34). Oxford University Press.

- Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference*, 15, 17–29.
- Chalmers, D. J. (2020). What is conceptual engineering and what should it be? *Inquiry*, 1–18. <https://doi.org/10.1080/0020174X.2020.1817141>
- Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition: The twentieth annual Carnegie symposium on cognition*. Erlbaum.
- Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58(1), 7–19. <https://doi.org/10.1111/1467-8284.00096>
- Dembroff, R. (2016). What is sexual orientation? *Philosophers' Imprint*, 16(3), 1–27. Permalink: <http://hdl.handle.net/2027/spo.3521354.0016.003>
- Deutsch, M. (2020). Speaker's reference, stipulation, and a dilemma for conceptual engineers. *Philosophical Studies*, 177, 3935–3957. <https://doi.org/10.1007/s11098-020-01416-z>
- Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology*, 41, 451–463. <https://doi.org/10.1037/0012-1649.41.3.451>
- Diesendruck, G., Goldfein-Elbaz, R., Rhodes, M., Gelman, S., & Neumark, N. (2013). Cross-cultural differences in children's beliefs about the objectivity of social categories. *Child Development*, 84(6), 1906–1917. <https://doi.org/10.1111/cdev.12108>
- Fassio, D., & McKenna, R. (2015). Revisionary epistemology. *Inquiry*, 58(7–8), 755–779. <https://doi.org/10.1080/0020174X.2015.1083468>
- Fischer, E. (2020). Conceptual control: On the feasibility of conceptual engineering. *Inquiry*, 1–29. <https://doi.org/10.1080/0020174X.2020.1773309>
- Flocke, V. (2021). How to engineer a concept. *Philosophical Studies*, 178, 3069–3083. <https://doi.org/10.1007/s11098-020-01570-4>
- Foster-Hanson, E., & Rhodes, M. (2020). The psychology of natural kind terms. In S. T. Biggs & H. Geirsson (Eds.), *The Routledge handbook on linguistic reference*. Routledge.
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. Oxford University Press.
- Gelman, S. A., & Coley, J. D. (1990). The importance of knowing a dodo is a bird: Categories and inference in 2-year-old children. *Developmental Psychology*, 26(5), 796–804. <https://doi.org/10.1037/0012-1649.26.5.796>
- Gelman, S. A., & Heyman, G. D. (1999). Carrot-eaters and creature-believers: The effects of lexicalization on children's inferences about social categories. *Psychological Science*, 10(6), 489–493. <https://doi.org/10.1111/1467-9280.00194>
- Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, 23, 183–209. [https://doi.org/10.1016/0010-0277\(86\)90034-x](https://doi.org/10.1016/0010-0277(86)90034-x)
- Golinkoff, R. M., Hirsh-Pasek, K., Bailey, L. M., & Wenger, N. R. (1992). Young children and adults use lexical principles to learn new nouns. *Developmental Psychology*, 28(1), 99–108. <https://doi.org/10.1037/0012-1649.28.1.99>
- Grassman, S., Schulze, C., & Tomasello, M. (2015). Children's level of word knowledge predicts their exclusion of familiar objects as referents of novel words. *Frontiers in Psychology*, 6, 1200. <https://doi.org/10.3389/fpsyg.2015.01200>
- Habgood-Coote, J. (2019). Stop talking about fake news! *Inquiry*, 62(9–10), 1033–1065. <https://doi.org/10.1080/0020174X.2018.1508363>
- Halberda, J. (2006). Is this a dax which I see before me? Use of the logical argument disjunctive syllogism supports word-learning in children and adults. *Cognitive Psychology*, 53, 310–344. <https://doi.org/10.1016/j.cogpsych.2006.04.003>
- Hall, D. G. (1991). Acquiring proper names for familiar and unfamiliar animate objects: Two-year-olds' word-learning biases. *Child Development*, 62(5), 1142–1154. <https://doi.org/10.2307/1131158>
- Haslanger, S. A. (2000). Gender and race: (What) are they? (what) do we want them to be? *Nous*, 34, 31–55. <https://doi.org/10.1111/0029-4624.00201>
- Haslanger, S. A. (2012). *Resisting reality: Social construction and social Critique*. Oxford University Press.
- Hutchinson, J. E. (1986). *Children's sensitivity to the contrastive use of object category terms*. Paper presented at the 1986 Stanford Child Language Research Forum, Stanford, CA.
- Isaac, M. G., Koch, S., & Neftci, R. (2022). Conceptual engineering: A roadmap to practice. *Philosophy Compass*, 17(10), 1–15. <https://doi.org/10.1111/phc3.12879>

- Kagan, J. (1981). *The second year*. Harvard University Press.
- Keleman, D., & Carey, S. (2007). The essence of artifacts: Developing the design stance. In E. Margolis & S. Lawrence (Eds.), *Creations of the mind: Theories of artifacts and their representation* (pp. 212–230). Oxford University Press.
- Knoll, V. (2021). Verbal disputes and topic continuity. *Inquiry*, 1–22. <https://doi.org/10.1080/0020174X.2020.1850340>
- Koch, S. (2023). Why conceptual engineers should not worry about topics. *Erkenntnis*, 88, 2123–2143. <https://doi.org/10.1007/s10670-021-00446-1>
- Kocurek, A. (2022). What topic continuity problem? *Inquiry*, 1–21. <https://doi.org/10.1080/0020174X.2022.2126884>
- Koslow, A. (2022). Meaning change and changing meaning. *Synthese*, 200(94). <https://doi.org/10.1007/s11229-022-03563-8>
- Kucker, S. C., McMurray, B., & Samuleson, L. K. (2018). Too much of a good thing: How novelty biases and vocabulary influence known and novel referent selection in 18-month-old children and associative learning models. *Cognitive Science*, 42(Suppl 2), 463–493. <https://doi.org/10.1111/cogs.12610>
- Landes, E. (forthcoming). How language teaches and misleads: “Coronavirus” and “Social Distancing” as case studies. In M. G. Isaac, S. Koch, & K. Scharp (Eds.), *New perspectives on conceptual engineering*. Synthese Library.
- Leslie, S. J. (2017). The original sin of cognition: Fear prejudice, and generalization. *Journal of Philosophy*, 114(8), 393–421. <https://doi.org/10.5840/jphil2017114828>
- Liittschwager, J. C., & Markman, E. M. (1994). 16-Month-old and 24-month-old—Use of mutual exclusivity as a default assumption in 2nd-label learning. *Developmental Psychology*, 30(6), 955–968. <https://doi.org/10.1037/0012-1649.30.6.955>
- Macnamara, J. (1982). *Names for things: A study of human learning*. MIT Press.
- Manne, K. (2017). *Down girl. The logic of misogyny*. Oxford University Press.
- Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.
- Markman, E. M. (1992). Constraints on word learning: Speculations about their nature, origins, and domain specificity. In M. R. Gunnar & M. P. Maratsos (Eds.), *Modularity and constraints on language and cognition: The Minnesota symposium on child psychology*. Erlbaum.
- Markman, E. M., & Wachtel, G. F. (1988). Children’s use of mutual exclusivity to constrain the meaning of words. *Cognitive Psychology*, 20, 121–157. [https://doi.org/10.1016/0010-0285\(88\)90017-5](https://doi.org/10.1016/0010-0285(88)90017-5)
- Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children’s word learning. *Monographs of the Society for Research in Child Development*, 54(3–4), 1–132. <https://doi.org/10.2307/1166130>
- Merriman, W. E., Evey-Burkey, J. A., Marazita, J. M., & Jarvis, L. H. (1996). Young two-year-olds’ tendency to map novel verbs onto novel actions. *Journal of Experimental Child Psychology*, 63, 466–498. <https://doi.org/10.1006/jecp.1996.0059>
- Mervis, C. B. (1987). Child-basic object categories and early lexical development. In U. Neisser (Ed.), *Concepts and conceptual development: Ecological and intellectual factors in categorization* (pp. 201–233). Cambridge University Press.
- Mervis, C. B., Golinkoff, R. M., & Bertrand, J. (1994). Two-year-olds readily learn multiple labels for the same basic-level kind. *Child Development*, 65, 1163–1177.
- Nado, J. (2023). Classification procedures as the targets of conceptual engineering. *Philosophy and Phenomenological Research*, 106(1), 136–156. <https://doi.org/10.1111/phpr.12843>
- Nimtz, C. (2021). Engineering concepts by engineering social norms: Solving the implementation challenge. *Inquiry*, 1–28. <https://doi.org/10.1080/0020174X.2021.1956368>
- Prasada, S., Hennefield, L., & Otap, D. (2012). Conceptual and linguistic representations of kinds and classes. *Cognitive Science*, 36(7), 1224–1250. <https://doi.org/10.1111/j.1551-6709.2012.01254.x>
- Quine, W. V. O. (1960). *Word and object*. MIT Press.
- Rhodes, M., & Gelman, S. A. (2009). A developmental examination of the conceptual structure of animal, artifact, and human social categories across two cultural contexts. *Cognitive Psychology*, 59(3), 244–274. <https://doi.org/10.1016/j.cogpsych.2009.05.001>
- Ritchie, K. (2021a). Essentializing language and the prospects for ameliorative projects. *Ethics*, 131(3), 460–488. <https://doi.org/10.1086/712576>

- Ritchie, K. (2021b). Essentializing inferences. *Mind & Language*, 36, 570–591. <https://doi.org/10.1111/mila.12360>
- Samuelson, L. K., & McMurray, B. (2017). What does it take to learn a word? *WIREs Cognitive Science*, 8(1–2), e1421. <https://doi.org/10.1002/wcs.1421>
- Schnell, L. (2020). Coronavirus and social distancing: Why people won't avoid each other. *USA Today*, March 17.
- Sterken, R. (2020). Linguistic intervention and transformative communicative disruptions. In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual engineering and conceptual ethics* (pp. 417–434). Oxford University Press.
- Strawson, P. F. (1963). Carnap's view on constructed systems versus natural languages in analytic philosophy: The two methods. In P. A. Schilpp (Ed.), *The philosophy of Rudolf Carnap* (pp. 502–519). Open Court.
- Szabó Gendler, T. (2008). Alief and belief. *Journal of Philosophy*, 105(10), 634–663.
- Thomasson, A. L. (2021). Conceptual engineering. When do we need it? How can we do it? *Inquiry*, 1–26. <https://doi.org/10.1080/0020174X.2021.2000118>
- Walton, G. M., & Banaji, M. R. (2004). Being what you say: The effect of essentialist linguistic labels on preferences. *Social Cognition*, 22, 193–213. <https://doi.org/10.1521/soco.22.2.193.35463>
- Wodak, D., Leslie, S.-J., & Rhodes, M. (2015). What a loaded generalization: Generics and social cognition. *Philosophy Compass*, 10(9), 625–635. <https://doi.org/10.1111/phc3.12250>

How to cite this article: Koch, S. (2023). How words matter: A psycholinguistic argument for meaning revision. *Mind & Language*, 1–17. <https://doi.org/10.1111/mila.12494>