In this paper, we argue that, in order to account for the apparently substantive nature of definitional disputes, a commitment to what we call 'Socratic essentialism' is needed. We defend Socratic essentialism against a prominent neo-Carnapian challenge according to which apparently substantive definitional disputes always in some way trace back to disagreements over how expressions belonging to a particular language or concepts belonging to a certain conceptual scheme are properly used. Socratic essentialism, we argue, is not threatened by the possibility that some apparently substantive definitional disputes may turn out to be verbal or conceptual, since this pluralist strategy, in our view, requires a commitment to more, rather than fewer, essences. What is more, a deflationary, metaphysically 'light-weight' construal of the essence-ascriptions in question leads to a peculiar conception of the pursuit of metaphysicians as behaving like deceptive (or self-deceived) grammarians pretending to be scientists. Moreover, this deflationary attitude, we argue, spreads beyond metaphysics and philosophy more broadly to apparently substantive definitional disputes in the sciences as well as other in other disciplines, such as art criticism.

**KEYWORDS**
definitional disputes, essentialism, neo-Carnapian approaches to metaphysics
1 | SOCRATIC ESSENTIALISM

On the face of it, such questions as ‘What are forces?’, ‘What is knowledge?’, and ‘What is money?’, bear on the nature or essence of the phenomenon in question, viz., forces, knowledge, and money. More generally, when philosophers and scientists carry on debates over the definition of F, where ‘F’ is a general term, they typically assume that their inquiries bear not (or not just) on the meaning of the word ‘F’ or the concept F. (In what follows, we use italics for concepts and quotation marks for linguistic expressions.) Thus, instead of asking ‘What is F?’ or ‘What is to be F?’, one may also ask ‘What is this thing called F?’ This formulation makes clear that the question is not about the terms or concepts we use to represent F; rather, definitional disputes aim at understanding what the phenomenon, F, itself is. When Socrates challenges Meno’s definition of virtue as being nothing else but the ability to rule (73d), on the grounds that people who rule unjustly are not virtuous, Socrates is not (or at least not only) claiming that the (Greek analogue ‘arête’ of the English) term ‘virtue’ fails to apply to unjust rulers; rather, he is claiming that virtue is not the ability to rule. In other words, Socrates and Meno disagree about the nature of virtue. Both assume that there is something that virtue is, that distinguishes it from other phenomena, and about which one can be right or wrong.

When we speak of the *essence* or *nature* of F (where ‘essence’ and ‘nature’ are used interchangeably), we mean standardly that in virtue of which F is what it is. That such questions as ‘What is piety?’ concern the essence or nature of the phenomenon under investigation functions as a dialectical presupposition in Plato’s early Socratic dialogues. This dialectical presupposition is assumed to be required in order for such Socratic questions to have correct or incorrect answers and to be meaningfully debatable. Essences, otherwise put, are what we are after when we ask Socratic questions of the form, ‘What is F?’.

In what follows, we shall focus on definitional questions of the form ‘What is F?’ or ‘What it is to be F?’, where ‘F’ is a general term. We take the answers to such definitional questions (insofar as they have answers) to be *real definitions* which purport to state the essence of a kind of phenomenon. Real definitions contrast with *nominal definitions* which state the meaning of an expression or a concept that is used to refer to the phenomenon in question. While the essential properties mentioned in a real definition are also generally taken to be necessary properties of the defined entity, the converse does not always hold: not every necessary property is also plausibly taken to be an essential property (Fine, 1994). To illustrate, being loved by the gods, as Socrates and Euthyphro discover over the course of Plato’s *Euthyphro*, turns out to be merely an accidental feature of what is pious, even though it is one that (or so we may suppose) applies necessarily to all and only what is pious. The fact that something is pious, as Socrates argues, explains the fact that it is loved by the gods, rather than the other way around.

Essences have traditionally been thought to play important explanatory roles: in particular, they are invoked to help answer central questions about existence, identity, persistence and modality (i.e., necessity or possibility). To illustrate, what determines when an entity comes into or goes out of existence? In virtue of what is an entity identical to itself at a time or over time? Through what sorts of changes can an entity persist? What explains what features it must have in order for it to exist at all, and which features are compatible with its nature but optional (i.e., its modal profile)? Unless responses to these central questions can be found, key questions about the world will go unanswered. Essentialists can appeal to essences in their response to these questions concerning existence, identity, persistence, and modality, while anti-essentialists must either reject the demands for an answer to these questions as somehow misplaced or formulate their responses by way of some essence-free analogue.

Socratic essentialism (SE), as we understand it, maintains that some (but not necessarily all) phenomena have essences, and that these essences can be rigorously studied by philosophical or scientific means. (SE) differs from other stronger formulations of essentialism in its fairly minimal commitments. On the one hand, (SE) does not require that all phenomena have essences; rather, (SE) is compatible with the idea that some phenomena (e.g., jade, the common cold, or grueness), which might initially appear to have essences, upon reflection turn out to be a heterogeneous

---

1 By contrast, we take definitional questions of the form ‘What is a?’ or ‘What is it to be a?’, where ‘a’ is a singular term, to concern individual essences. These questions, though of course a worthwhile topic in their own right, are not our direct focus here.
collection of distinct phenomena to which no unified essence can be ascribed. On the other hand, (SE) is also not committed to the view that essences across the board must be modeled along the lines of Kripke/Putnam-style natural kind essences, as (i) intrinsic or microstructural (e.g., physical, chemical, or biological); (ii) hidden or discoverable only by a posteriori means; and (iii) independent of human categorizations or other activities. Rather, (SE) allows that the essences of some phenomena are not only extrinsic (e.g., historical), but also dependent in certain ways on contingent human factors (e.g., in the case of games, on the conventional rules governing the practice).

Grasping the nature or essence of the subject-matter that is being investigated is crucial for our understanding of reality and calls for the formulation of real definitions whose job it is to state the nature or essence of the phenomena in question. Yet essentialism has been subjected to serious criticisms coming from many different corners. For one thing, inspired by sceptically minded early modern philosophers and scientists (e.g., Bacon, Hobbes, and, to some extent, Locke), modern-day anti-essentialists harbour doubts concerning both the existence of essences and our ability to grasp them, and interpret apparent essence-ascriptions instead as in some way reflective of our specifically human categorizations or other practices, rather than of the world ‘as such’ (Quine, 1953). Secondly, essentialism has come under attack by scientists and philosophers of science who argue that this doctrine is incompatible with both Darwinian evolutionary theory and the pluralist ‘turn to practice’ approach which emphasizes the multi-faceted goals and interests driving actual scientific practice (Dupré, 1993; Ereshefsky, 2004; Waters, 2017). While we consider these challenges to be interesting and important in their own right, our present goal is to address a different challenge to essentialism, which has also recently risen to prominence in the wake of the popularity of neo-Carnapian deflationist approaches to meta-metaphysics. According to this third challenge, a commitment to essentialism cannot be justified by appeal to the apparently substantive nature of definitional disputes, since the disputes in question can, and should, be construed as being merely verbal.

Before turning to the verbal dispute challenge to essentialism, two further points about (SE) should be noted. First, despite the serious concerns just cited which have been raised against essentialism, we are struck by the fact that the main reasons in favor of essentialism are in fact left intact even in the face of these challenges. First, for those who are engaged in the study of a particular natural or social phenomenon, the question of what this phenomenon is, i.e., its nature, remains of central interest. Such questions as ‘What is money?’ for example, have not lost their appeal despite concerns over the application of essentialism to social reality. One reason for this is that if one aims at elucidating, e.g., the origin, effects or function of a phenomenon, the question of what its nature is has to be at least tacitly answered. Relatedly, disagreements concerning the correct answers to these questions typically appear to presuppose essences. To carry on a meaningful debate concerning money, for example, one has to agree that there is a common subject-matter, viz., money, whose nature can be disputed by competing theorists. Second, the need to disentangle the essential from the accidental features of natural or social phenomena remains a significant concern. Money is said to be a means of exchange, a measure of value, a store of value, a symbol of success, and a means of influence. But which of these features, if any, are essential to money? Which are merely accidental? Arguably, such questions must be addressed if we want to make progress in our scientific understanding of natural and social reality. Third, one main motivation for the search of definitions is the desire to avoid misunderstanding: attempts to define a natural or social phenomenon are useful in unifying research areas whose practitioners misunderstand, talk past, or ignore each other, because they lack a common vocabulary. Shared definitions and an agreed-upon subject-matter are a key to scientific progress. The fact that objections raised against essentialism have left many of these motivations in favor of essentialism untouched suggests that there is room for an essentialist framework, provided that it can meet prominent anti-essentialist challenges.

2 | THE VERBAL DISPUTE CHALLENGE TO SOCRATIC ESSENTIALISM

A new type of objection to essentialism has recently emerged which throws doubt on the Socratic idea that essences are presupposed by definitional disputes. According to this objection, most if not all definitional disputes
turn out to be merely verbal, rather than substantive, so that postulating essences is not needed to make sense of them. Three considerations have been invoked in favor of this deflationist conception of definitional disputes.

First, definitional disagreements in philosophy persist without apparently ever being resolved, which suggests that, in pursuing these disputes in perpetuity, we fail to make progress in our grasp of underlying essences. Socratic dialogues investigate the nature of pleasure, virtue, knowledge, love, wisdom, piety, courage, friendship (cf., Plato’s *Philebus, Meno, Theaetetus, Symposium, Charmides, Euthyphro, Laches* and *Lysis*, respectively). Over two-thousand years later, the question of what the nature of these phenomena is (or whether they even have a nature) remains highly controversial among philosophers, thereby calling into question the fruitfulness of a Socratic investigation into the essences of things. Chalmers (2015), for example, argues that many disagreements in philosophy, in particular definitional disagreements, persist in this way: definitions are rarely agreed upon in philosophy.

Second, pluralist answers to ‘What is F?’ questions have become increasingly popular in various philosophical fields. By ‘pluralism’ we mean the view according to which what initially appears to be a phenomenon of a single unified type, upon scrutiny, turns out to consist of multiple phenomena of distinct types. To illustrate, truth, knowledge, responsibility, or justification, according to the pluralist, can turn out to be not one phenomenon, but a variety of distinct phenomena. Pluralistic positions (also known as polysemy, polyadic, disjunctive, multi-dimensional, or cluster theories) have been defended, for example, about truth (Pedersen & Wright, 2018); justification (BonJour, 2002; Goldman, 1988); pain (Borg et al., 2020; Corns, 2014, 2020); as well as senses (Macpherson, 2011) or sounds (Killin, 2022). Chalmers (2011) suggests that ‘pluralism should be the default view for almost any philosophical concept’ (2011, p. 540). As pluralistic approaches to various topics gain traction, the project of discovering the essence of the phenomenon F in question loses its luster. If, each time we set out to uncover the nature of F, we end up hitting upon distinct varieties of F that have little in common, we should arguably shelve the project of identifying real definitions for the phenomena at issue.

Third, progress in our understanding and diagnosis of verbal disagreements has rendered the view more plausible that such disagreements are indeed widespread. Although the nature of verbal disagreement itself has become a matter of dispute (see, e.g., Belleri, 2018; Chalmers, 2011; Hirsch, 2009; Jenkins, 2014; Thomasson, 2017b; Vermeulen, 2018), we do seem to have a workable pre-theoretical grasp of the distinction between substantive and verbal disagreements. Socrates’ and Meno’s disagreement about virtue, for example, appears to be substantive: they disagree about the nature of virtue. By contrast, suppose that, in the course of a discussion of the definition of virtue, one of those participating in the dispute puts forward the proposal that a virtue is an angel responsible for supervising the movements of the heavenly bodies. Interpretive charity would lead us to infer that this person means something quite different by ‘virtue’ from what the rest of us have in mind when we use this term. (In the Christian angelic hierarchy, a virtue is a type of angel.) Entering into a dispute over the nature of virtue with this interlocutor, while failing to notice that the participants in the dispute are employing the term ‘virtue’ in a different sense, would result in a purely verbal disagreement. Indeed, in the case in question, the disputants may in fact agree with one another over what each of them calls ‘virtue’.

We shall here, following Chalmers (2011), assume that a verbal disagreement is a disagreement whose sole source lies in different meanings ascribed to a single word. (The details of Chalmers’ account of verbal disagreements, however, are not crucial either to the characterization of the objection or to the response we go on to discuss.) In order to have a substantive definitional disagreement, therefore, it is necessary to clear away all possible misunderstandings about the meaning of the words that are being used in stating the disagreement in question. If, after this clarification has been achieved, the disagreement persists, then we can safely conclude that the disagreement in question concerns the nature of the phenomenon under discussion. Chalmers (2011) proposes a simple method for diagnosing verbal definitional disagreement along these lines: introduce subscripts ‘F_1’ and ‘F_2’, and bar the use of the term ‘F’. If, after this strategy has been implemented, the disputing parties no longer disagree about F_1 and F_2 (or at least not in a way that is relevant to the original dispute), then the original definitional dispute was verbal.
Philosophers like to think that the majority of definitional disputes are substantive, while only a minority are verbal. Chalmers (2011) maintains that the opposite is the case. More often than not, he claims, seemingly substantive disagreements turn out to be merely verbal. Chalmers concludes that, instead of focusing on ‘What is \( F \)?’ questions, which are all too often purely verbal, we should instead investigate the function that \( F \) could and should play:

This picture leads to a certain deflationism about […] the interest of questions such as ‘What is \( [F] \)?’ or ‘What is it to be \( [F] \)?’ Some component of these questions is inevitably verbal, and the nonverbal residue can be found without using ‘\( [F] \)’.

On the picture I favor, instead of asking ‘What is \( [F] \)?,’ one should focus on the roles one wants \( [F] \) to play and see what can play that role. The roles in question here may in principle be properties of all sorts: so one focuses on the properties one wants \( [F] \) to have and figures out what has those properties. But very frequently, they will be causal roles, normative roles, and especially explanatory roles.

For example […] instead of asking ‘What is a belief? What is it to believe?’ and expecting a determinate answer, one can instead focus on the various roles one wants belief to play and say: here are some interesting states: \( B_1 \) can play these roles, \( B_2 \) can play these roles, \( B_3 \) can play these roles. Not much hangs on the residual verbal question of which is really belief. (Chalmers, 2011, p. 538; see also Chalmers, 2020)

Chalmers considers verbal disputes to be of little or no value: all verbal disputes should be avoided and once a verbal dispute about \( F \) has been diagnosed, the question of what ‘\( F \)’ really means is of little importance outside the study of this peculiar form of expression. But one may also endorse the ubiquity of verbal disputes and yet still consider them to be of some value. Cappelen, while agreeing with Chalmers that definitional verbal disputes are ubiquitous (Cappelen, 2017, p. 59), disagrees with the sentiment concerning the value of verbal disputes Chalmers voices in the passage cited above (Cappelen, 2018, pp. 191ff; 2020). According to Cappelen, even if it is pointless to carry on a dispute concerning the nature of \( F \) while disagreeing over the meaning of ‘\( F \)’, the question of whether ‘\( F \)’ refers to \( F_1 \) or \( F_2 \) does matter. It is even, Cappelen argues, presupposed by the conceptual pluralism advocated by Chalmers. For there must be, in Cappelen’s view, some reason why \( B_1 \) and \( B_2 \) are associated with ‘belief’ and not with, e.g., ‘Apple’ or ‘Sensation’. The reason in question pertains to ‘Belief’, simpliciter: the term is an ‘anchor point’ which we cannot drop on pain of lacking any ground to include some properties but not others in the belief cluster.

Chalmers and Cappelen nonetheless agree that verbal disputes are not genuine disagreements: if disputing parties use the same word with different meanings, they do not in fact disagree. This view is rejected by Plunkett and Sundell (2013), Plunkett (2015) and Thomasson (2017a, 2020). According to them, one important and genuine kind of disagreement is metalinguistic and can occur even when speakers use the same words with different meanings. Suppose A and B carry on a dispute over the question of whether a particular curry dish is spicy, while using the term ‘spicy’ to denote different levels of spiciness. In this case, their dispute may bear not on the spiciness of the curry (assuming, for example, that the curry tastes exactly the same for A and B), but on how the term ‘spicy’ should be used. The reason why such disputes are possible is that, in addition to the contents that are semantically expressed by one’s statements, there are other contents which are ‘communicated pragmatically and not semantically’ (Plunkett & Sundell, 2013, p. 13): these pragmatically conveyed contents –here, about the proper use of the term ‘spicy’– can be the locus of disagreements, when disputants mean different things by the words they use.
Such a ‘metalinguistic negotiation’ strategy is especially relevant to definitional disputes that are of concern to our present discussion. Suppose that Mary holds that to be happy is to have a positive balance of pleasure over pain; Bob counters that to be happy is to be satisfied with one’s life. While Mary and Bob may seem to disagree about the nature of happiness, an admittedly tempting deflationist take on their dispute is that it is in fact a metalinguistic disagreement over how the word ‘happiness’ should be used. Likewise, Thomasson (2017a) considers disagreements over the definition of art and argues that many such disagreements about what art is can be read as metalinguistic negotiations over how properly to employ the term ‘art’ or the concept, art. (We will return to this point in more detail in Section 6 below.)

Such debates about the value of verbal disputes are orthogonal to the deflationist argument which concerns us here. Whether verbal disputes are only superficially interesting or not, and whether (many) verbal disputes constitute genuine disagreements or not, the point remains that if definitional disputes in metaphysics are verbal, then the essentialist approach to definitional inquiries appears to be in danger. For it then seems to be no longer the case that definitional disputes presuppose essences. Instead of asking ‘What is F?’ to discover F’s nature, deflationists argue that we should give priority to question such as ‘What does ‘F’ mean?, ‘What should ‘F’ mean?, ‘Which roles could or should the term ‘F’ or the concept F play?’, ‘What should the extension and intension of the concept F be?’, ‘What should count as an F?’ (or perhaps ‘What should F be?’, in the objectual but metaphysically lightweight understanding of Cappelen, 2018, Ch. 12).

Summing up, while definitional inquiries of the form ‘What is F?’ understood as inquiries about the nature of F, have long been optimistically pursued by philosophers, three converging considerations suggest that such disputes are often, if not always, verbal: (i) little or no agreement is ever reached concerning the nature of F; (ii) philosophers in various fields have independently arrived at pluralistic conclusions to the effect that, for many Fs, there is no unique thing that F is; and (iii), by barring talk of ‘F’, many such disagreements vanish. On the basis of these three points, it is concluded that inquiries into the essence of F should be given up, and one should instead focus on the role or function that the concept F or the word ‘F’ plays or ought to play.

3 | ANSWERING THE VERBAL DISPUTE CHALLENGE TO SOCRATIC ESSENTIALISM

One answer to the verbal dispute challenge is that, so long as some definitional disputes remain non-verbal, essences are still required to make sense of these. We agree, of course, that if some definitional disputes are non-verbal, the Socratic essentialist should maintain that these disputes concern the essences of the phenomena under consideration. But that is a big ‘if’. For if Chalmers (2011, p. 532n) is right that ‘almost all’ definitional questions in philosophy are ‘beset by verbal disputes, in a fashion that is occasionally but too rarely recognized’, then Socratic essentialism is limited to a narrow and shrinkingly small set of definitional disputes. We shall now argue, by contrast, that, even if verbal disputes are ubiquitous, this outcome is in fact favorable to the Socratic essentialist, since we then have reason to recognize more, rather than fewer, essences than it at first appeared. On this view, the more disputes are characterized as verbal, the better for (SE).

Why should the existence of definitional disputes provide evidence in favor of, rather than against, (SE)? Our proposal is that, in order for a disagreement about the nature of F to turn out to be verbal, one needs to have reached a substantive agreement to the effect that F₁ and F₂ have distinct natures. Suppose that Mary and Bob disagree about the nature of shame: Mary thinks that shame can only be directed at oneself, while Bob thinks that one can also be ashamed of one’s actions. After heated debates, Mary and Bob come to realize that they were talking past each other: there are in fact two different emotions in the vicinity. Mary was speaking about what is commonly called ‘shame’; by contrast, Bob had in mind the emotion that is also known as ‘guilt’. Mary and Bob may now agree or not agree to employ the terms ‘shame’ and ‘guilt’ in the way just mentioned. Regardless of whether they both sign on to this way of using the terms in question, however, they have now come to agree that their
original dispute concerned two distinct emotions: one, ordinarily called ‘shame’ which is primarily directed at one-
self; and another, ordinarily called ‘guilt’, which is primarily directed at one’s behaviour (see Deonna et al., 2012;
Lewis, 1971; Mulligan, 2009, for accounts of the shame/guilt distinction along these lines). Instead of dealing with
just one phenomenon, whose essence is in dispute, we now have two phenomena, which are agreed to be distinct.
In the course of inquiring into the nature of F (shame), we have discovered a distinction between two new phe-
nomena: F\textsubscript{1} (the denotation\textsuperscript{2} of ‘shame’, used in the ordinary sense) and F\textsubscript{2} (the denotation of ‘guilt’, used in the
ordinary sense), a previously unnoticed distinction which, according to (SE), itself can be traced to distinct es-
sences. Though the question ‘What is F?’ has turned out to be verbal, the reason why it has turned out to be so is
that a more fundamental agreement has been reached to the effect that there is a non-verbal distinction between
F\textsubscript{1} and F\textsubscript{2}.

In addition to reaching an agreement concerning the distinction between F\textsubscript{1} and F\textsubscript{2}, a whole new area of
metaphysical inquiry has thereby opened up: the questions ‘What is F\textsubscript{1}?’, ‘What is F\textsubscript{2}?’; and ‘How are F\textsubscript{1} and
F\textsubscript{2} related?’ can now be placed on our agenda. If this is correct, the upholder of the verbal dispute objection
against essentialism sounds like someone who, when hunting for jade, discovers sparkling pieces of jadeite
and nephrite, but leaves them by the wayside in order to dedicate his life to improving the concept, jaded.
Diagnosing a verbal dispute by no means forces us to give up on our essentialist commitments in order to em-
brace a new kind of inquiry into ‘the roles one wants [F] to play’, as Chalmers puts it in the passage cited above.
On the contrary, we now face an even more ambitious and promising inquiry into the essences of things. This
inquiry includes making clear the key differences between F\textsubscript{1} and F\textsubscript{2}, as well as elucidating their connections,
which may in turn explain why the two phenomena were conflated to begin with and why a single term or
concept was used for both. Indeed, once the distinctions and relations between the various phenomena picked
out by a polysemic expression have been clarified, the recognition that a transfer of meaning has taken place
often provides a natural explanation of the polysemy of the expressions in question (see Nunberg (1995);
Langacker (1993); Kövecses and Radden (1998); Liebesman and Magidor (2018), for important developments
of this idea).

Thus, what the ubiquity of definitional verbal disputes shows, for the upholder of (SE), is not that there are
fewer essences than expected, but that there are more. In addition, it turns out that there was in fact more
agreement about essences than previously thought. But what if the question ‘What is F\textsubscript{1}?’ now turns out to be
verbal as well? What if ‘F\textsubscript{1}’ (e.g., ‘guilt’) turns out to denote two distinct types of phenomena, ‘F\textsubscript{1}\textsuperscript{1}’ (e.g., guilt, in
the strict sense) and ‘F\textsubscript{1}\textsuperscript{2}’ (e.g., remorse)? In that case, we have reached an even more basic agreement about
the distinction between ‘F\textsubscript{1}\textsuperscript{1}’ and ‘F\textsubscript{1}\textsuperscript{2}’. On the picture we propose, one key area of progress in essentialist meta-
physics thus consists precisely in drawing ever more fine-grained essential distinctions between types of en-
tities, and discerning essential relations between phenomena either in the sense of relations holding between
these entities in virtue of their essences or in the sense of relations holding between these and other entities,
when the essences in question are extrinsic. The picture that emerges is one of an ever-growing network of
interrelated essences (see, e.g., Strawson, 1992, Ch. 2, for a similar point about concepts). On this picture, es-
sentialist inquiries typically progress by discovering overlooked essential distinctions or connections between
different types of phenomena.

4 | THREE EXAMPLES

Let us apply the proposed essentialist framework to three definitional disputes in philosophy. Consider, first, the
following disagreement concerning forgiveness. Suppose that A thinks that forgiving is a speech act, which needs
to be heard by the person one forgives, while B thinks instead that forgiving is an internal episode. B argues that

\textsuperscript{2}We use ‘denotation’ in a broad way, to mean the semantic value of a term.
one can utter the phrase 'I forgive you', and never really forgive the person. By contrast, A argues that ceasing to resent somebody without ever taking the step of making clear to her that we have forgiven her is not enough to forgive. After discussing the matter, A and B come to agree that one should distinguish forgiving as a mental episode from forgiving as a speech act that needs to be heard by the person one forgives (see Milan, 2022, for references and discussion on this distinction; similar distinctions arise for judging, blaming, committing, or consenting). This opens up the question of how these phenomena are related: one answer to this question is that the speech act of forgiving expresses the mental state of forgiving, in the same way as the speech act of ordering x to F expresses the speaker’s intention that x Fs.

Consider, next, the question ‘What is a sound?’. Suppose that A maintains that sounds are waves in some medium, while B maintains that sounds are vibratory events internal to an object. Clearly, there is a distinction between two phenomena here, and the question of whether ‘sound’ refers to air waves or to an event internal to an object appears to be verbal. Since the distinction is accepted by both parties, it is tempting to conclude that the term, ‘sound’, is polysemic (Killin, 2022), in which case the dispute in question would be verbal. But the distinction between sounds as events internal to objects and sounds as waves is a substantive one, as is the question of how the two phenomena are related. One answer to this latter question is that sound waves transmit sound events (Casati & Dokic, 1994; O’Callaghan, 2007).

As a third and final example, consider the dispute over whether pleasures are sensations located in the body (e.g., Stumpf, 1928) or intentional states (e.g., Brentano, 2009, p. 113) directed at what one takes pleasure in. Participants in this debate agree that bodily sensations are not intentional states. To bracket the question of which of these phenomena deserves to be called ‘pleasure’, let us follow Feldman (2004) and call the former ‘sensory pleasures’ and the latter ‘attitudinal pleasures’. (Chalmers’ subscript strategy is also often adopted in the pain literature to distinguish pain, construed as an experience, from pain, construed as a bodily quality.) We can then move on to the question of what relates these two phenomena to one another. One proposal is that sensory pleasures are intentional objects of attitudinal pleasures: one takes pleasure in one’s sensory pleasures (Feldman, 2004).

5 | THE MODAL NORMATIVIST’S RESPONSE

So far, we have argued only that the classification of some apparently substantive definitional disputes as verbal disputes does not threaten (SE). But this does not yet show that an essentialist construal of such verbal disputes is in fact preferable to that offered by neo-Carnapians who favor a more deflationary construal of such disputes. In this and the following sections, we defend the claim that the essentialist construal of verbal disputes should in fact be preferred to that offered by deflationist metaphysicians. For the sake of concreteness, we focus on the modal normativist approach developed by Amie Thomasson, as a representative of the broader neo-Carnapian perspective (cf., especially, Thomasson, 2007, 2013, 2017a, 2017b, 2020, 2022).

Consider, again, the dispute between A and B described above concerning the nature of forgiveness. When faced with a dispute of this kind, (SE) maintains that the distinction in question between ‘forgiving'₁ (F₁) and ‘forgiving'₂ (F₂) is not just an actual distinction, but an essential distinction, i.e., that each of F₁ and F₂ has an essence associated with it, and that the essences in question are distinct. Thus, according to (SE), transitions like those in (1)–(6) from discerning an actual distinction to discerning an essential distinction are licensed:

1. Forgiving₁ is a speech act.
2. Forgiving₁ is essentially a speech act.
3. Forgiving₂ is an internal episode.
4. Forgiving₂ is essentially an internal episode.
5. Speech acts are distinct from internal episodes.
6. Speech acts are essentially distinct from internal episodes.
Given Thomasson’s approach, a modal normativist can in fact go along with accepting these transitions; however, in doing so, as Thomasson would maintain, the modal normativist does not incur a commitment to a ‘heavy-weight’ realist conception of essence or metaphysical modality.\(^3\) Rather, statements concerning essences, such as those in (2), (4), and (6), so Thomasson proposes, should be understood as ways of pragmatically conveying semantic rules regarding the use of object-language expressions in the object-language (Thomasson, 2007, pp. 143–144; Thomasson, 2020). Thus, if the transition from (1) to (2) is licensed, it would be licensed by the fact that the term, ‘forgiving', (i.e., the term ‘forgiving’ as it is used by A), is governed by a semantic rule to the effect of ‘Apply ‘forgiving’ only when ‘speech act’ applies'; and similarly for the transitions from (3) to (4) and from (5) to (6).

Once A and B have agreed to recognize a distinction between the terms, ‘F\(_1\)’ and ‘F\(_2\)’, they are no longer engaged in a verbal dispute over the meanings of these terms or over the semantic rules governing their application. Suppose, however, that A and B are metaphysicians who do not, or at least not explicitly, conceive of themselves as being engaged in a dispute over how properly to use the English expression, ‘forgiving’ (or an analogous expression in another language that would be translated into English as ‘forgiving’) or over how properly to employ the concept, forgiving, which is part of the conceptual schemes of twenty-first-century inhabitants of certain communities. Nor, or so we may assume, do A and B conceive of themselves as putting forward a particular practical conception of how one ought to live one’s life, especially as it pertains to one’s interactions with people towards whom one harbours resentment. Rather, A and B conceive of themselves as being engaged in a theoretical dispute over what it is to forgive someone.

How, in this case, would the modal normativist characterize what is going on in the dispute at hand between A and B? According to Thomasson, metaphysicians (like other competent speakers of a language) must of course have mastered the rules of the language, in the sense of being able to apply or refuse the application of expressions of the language properly in accordance with the rules governing their use. Like grammarians (but unlike most competent speakers), however, metaphysicians also have an explicit understanding of these rules, though metaphysicians convey the rules governing object-language expressions in the object-language, whereas grammarians state these rules governing object-language expressions in the meta-language (Thomasson, 2007, p. 151). In addition to conveying rules governing the use of individual terms (e.g., ‘freedom’), metaphysicians may also be interested in the consequences of these rules governing individual terms and in the connections between these rules and the rules governing other related terms (e.g., ‘choice’, ‘responsibility’, ‘blame’, ‘resentment’, etc.). This latter practice can sometimes lead to the discovery of surprising hidden inconsistencies in the rules of use governing the expressions of our language (Thomasson, 2007, p. 154). In such a case, metaphysicians may then propose (more or less conservative) revisions in the rules governing the terms of our language in order to repair the faulty usage (Thomasson, 2007, p. 155).

But why would it be beneficial from the point of view of a metaphysician to convey rules governing object-language expressions in the object-language, rather than to follow the grammarian’s practice of stating these rules explicitly by means of meta-language statements (Thomasson, 2007, p. 152)? For Thomasson, the pursuit of the ‘serious ontologist’ is to communicate pragmatically normative claims concerning our linguistic or conceptual scheme, rather than to report discoveries about the world. For example, when viewed from this perspective, a metaphysician who utters the statement, ‘Tables don’t exist’, might be interpreted as intending to implicate that there is something defective about how the term ‘table’ is standardly used and as intending that listeners be able to work out the implication in question. Metaphysicians who appear to be engaged in substantive definitional disputes, in Thomasson’s view, thus ultimately aim at advocating for a change in our linguistic practices or in our conceptual scheme (Thomasson, 2017a, pp. 22-23). But if their goal is to modify

---

\(^3\) Although Thomasson’s approach is couched in terms of statements concerning metaphysical modality (i.e., metaphysical necessity and possibility), rather than in terms of statements concerning essences, we will assume in what follows that Thomasson would apply the same modal normativist treatment to both cases.
the norms governing the use of a term, a more effective way of reaching their goal, so Thomasson proposes, may be to represent themselves as making a quasi-scientific discovery about the world (e.g., the real nature of persons or art), rather than to represent themselves as making recommendations about how a word or concept ought to be used. For if they made the true character of their goals explicit, their recommendations might then be ignored or fail to achieve uptake (Thomasson, 2017a, p. 25). In addition, ‘serious ontologists’ may also be self-deceived about what they are up to. If they think of themselves as being engaged in a quasi-scientific pursuit leading potentially to discoveries about the world, their work may appear to be more authoritative, objective, or respectable than it otherwise would, given its alleged affinity to the sciences (Thomasson, 2017a, pp. 25–27). Those metaphysicians who are not persuaded by the preceding characterization of the activity in which they are engaged, so Thomasson proposes, may instead take her to be putting forward a recommendation of what they ought to be doing (Thomasson, 2017a, p. 27).

6 | DEFINITIONAL DISPUTES IN SCIENCE AND ART CRITICISM

In the previous section, we outlined how, given Thomasson’s characterization, the work of the ‘serious ontologist’ turns out to be that of a deceptive (or possibly self-deceived) grammarian, pretending to be a scientist. Thomasson’s diagnosis of the pretence and (self)deceptiveness that is involved in putting forward statements of metaphysical modality might at least at first sight appear to be confined to the domain of metaphysics. Modal normativism, after all, is an account of metaphysical modality, and metaphysical modality, or so one might think, is of particular interest only to metaphysicians. This, however, is not the case: statements concerning metaphysical modality are also of significant interest outside of metaphysics as well as outside of philosophy more broadly. Thomasson’s revisionary diagnosis therefore can be seen to generalize beyond its initial more narrowly defined target. One chief reason why modal metaphysical claims are debated outside of philosophy stems from the fact that definitional endeavors occupy an important role in a wide range of fields.

To illustrate, consider the marginalist revolution that took place in economics at the end of the nineteenth century: Menger (1871), Jevons (1871) and Walras (1874) discovered nearly simultaneously, and independently of each another, the marginal theory of economic value, which led to the rejection of the labour theory of value espoused among others by Smith (1776), Ricardo (1817) and Marx (1867). Menger, Jevons and Walras were all convinced that they had a better theory of the nature of (exchange) value than that proposed by labour theorists. Essentialist talk is ubiquitous in the writings of Menger and Jevons, who speak repeatedly of the nature and essential qualities of value. Walras does not explicitly appeal to the essence of value, but insists that his pure political economy, unlike the applied conception, deals with abstract or ideal types (Walras, 1874: sixth lesson). All of these theorists share the conviction that economic essences or abstract types give rise to strongly necessary economic laws.

A normativist reading of the marginalist revolution as bearing ultimately on semantic rules for the use of ‘exchange value’ strikes us as highly revisionary when viewed against these authors’ own conception of their discoveries. This revisionary construal moreover also presents a striking departure from the attitude taken by economists at large towards the marginalist revolution. By contrast, the natural and standard take is that Menger, Jevons and Walras made a substantive discovery about the world, viz., about the nature of exchange value. In addition, the discovery in question is plausibly understood to be not of an empirical, but of an a priori nature, even though it was of course prompted by the empirical observation that exchanges are commonplace.

But how can we discover by a priori means what economic value is, so the modal normativist might ask? There appears to be nothing particularly mysterious about our epistemic access to the phenomena in question. The previous account of the nature of value (i.e., the labour theory) was beset by paradoxes. For example, some goods and services demand a lot of work in order to be produced, and yet have very low exchange value, e.g., the works of a dedicated artist who lacks talent. By contrast, some goods and services have high exchange value and yet
require little labour, e.g., an autograph by Queen Elizabeth. This apparent mismatch between the amount of labour required to produce something and the value that is ascribed to it in economic transactions prompted the idea that an alternative account of exchange value was needed. In the course of attempting to repair the deficiencies of the labour theory of value, the marginalist conception of the nature of value was discovered. Once this alternative solution to the difficulties in question is appreciated, the marginalist conception simultaneously fixes the problems to which the labour theory gave rise, while at the same time offering a satisfying account of the reasons motivating agents to engage in economic exchanges in the first place.

Admittedly, a single illustration will not put to rest the more general worries one may have concerning our epistemic access to essences. Among other things, it is still a disputed matter whether these epistemic questions are at best susceptible only to a piecemeal approach. Nevertheless, epistemic worries concerning the possibility of grasping essences do seem to lose their grip in a case in which those engaged in the dispute in question appear to understand the competing accounts of the nature of value that are on offer and acknowledge that one is superior to the other. Hence, we see little reason to give up the standard essentialist reading of this particular episode in the history of science. By contrast, the proposal that the marginalist revolution should be construed as the outcome of a successful metalinguistic negotiation, stated in object language, about how the expression ‘value’ or the concept value ought to be employed in economics is strongly at odds with a natural construal of the debate at issue and would bring with it a radical shift in the standard reading of this central episode of the history of economics.

In addition, the metalinguistic negotiation strategy has the disadvantage of being only theoretical and thereby disengaged from a commitment about the worldly phenomenon which motivates an agent’s behaviour in the first place, in the following sense. While economists can certainly utilize theoretical expressions and concepts, such as those connected with the phenomenon of economic value, in order to understand, explain, or predict an agent’s behaviour, these idioms themselves (or the linguistic or conceptual roles played by them) cannot in general be identified with the worldly phenomenon that motivates agents to engage in economic exchanges: for most economic agents do not entertain expressions or concepts that are construed along marginalist lines; and therefore their behaviour cannot be seen as being driven by the theoretical roles played by these expressions or concepts within the marginalist framework. By contrast, Socratic essentialists have the option of appealing to ‘heavy-weight’ essences as the real world explanans which itself partially motivates an agent’s economic behaviour: it is because of the nature of economic value, so (SE) would maintain, that agents engage in economic exchanges in the way that they do.4

Perhaps, Thomasson would be happy to accept that pretence and (self-)deception is also widespread within a priori economics, since theoreticians working in this domain after all have been historically influenced by essentialist metaphysicians (Smith, 1990). As long as genuinely empirical sciences, such as biology, are free from essentialist claims, so Thomasson might reason, her diagnosis of what goes on in apparently substantive definitional disputes can still be limited to pursuits which are either overtly, or implicitly, metaphysical. Indeed, explicit essentialist and modal talk is less commonplace in biology and the idea that biological species, for example, have essences remains a minority view, though it has recently garnered some important defenders (see, e.g., Devitt, 2008). However, less controversial examples of essentialist claims can be found at the very heart of biology. An illustration of such a statement is, for example, the claim that, among biological phenomena, there is a key distinction between organisms (e.g., cells), on the one hand, and processes (e.g., cell divisions), on the other hand. This essential distinction in turn entails further statements concerning metaphysical modality, such as the claim that necessarily no cell division is a cell. Such essentialist and modal claims are so uncontroversial and widely held that they are not usually considered to be even worth stating. Nevertheless, claims of this sort are of central importance and doing away with them would have dramatic effects within the domain of biology. These claims, furthermore, concern neither the proper use of words nor that of concepts: rather, the distinction in question between organisms and processes is understood as a distinction in rebus, not as a

4We are grateful to Amie Thomasson for helpful feedback in connection with this and other points raised in this paper.
KOSLICKI and MASSIN

The distinction concerning the rules of use governing the terms ‘organism’ and ‘process’ or the concepts associated with them. Thus, if Thomasson is right, then the pretence and (self)deception she takes to be characteristic of the ‘serious ontologist’ is in fact much more widespread and is not confined only to metaphysicians or those resembling them (e.g., a priori economists).

Finally, statements concerning metaphysical modality also play a central role outside of both metaphysics and science. An illustrative example can be taken from Thomasson’s own discussion of the famous art critic, Clive Bell (1914), who was interested in formulating a definition of visual art that would not exclude the painters of his time. Bell proposed that all visual art can be defined in terms of what he calls ‘significant form’. Thomasson offers the following reading of Bell’s proposal:

When Clive Bell (1914) argued that art is significant form, he may quite readily be seen not as just engaged in a shallow verbal dispute with those who would use ‘art’ to apply only to mimetic work, or as saying something that would have been trivially false, at least in an older art-historical context. Instead, regardless of how we interpret the literal content of his words, he can be seen as doing something non-trivial and reasonable: as pragmatically pressing for a new understanding of ‘art’ that could be more inclusive, and include the works of Cézanne, Gauguin and Matisse as much as of DaVinci and Tintoretto. (Thomasson, 2017a, p. 15)

If we take Bell at his own word, however, we see that Thomasson’s reading does not do justice to how Bell himself conceives of his own project. For Bell starts his essay by noting that ‘[e]veryone in his heart believes that there is a real distinction between works of art and all other objects; this belief my hypothesis justifies’ (Bell, 1914, p. 1). He then sets out to delineate the ‘nature of art’, ‘the essential quality in a work of art, the quality that distinguishes works of art from all other classes of objects’ and presents his proposed account as follows:

What is this quality? What quality is shared by all objects that provoke our aesthetic emotions? What quality is common to Sta. Sophia and the windows at Chartres, Mexican sculpture, a Persian bowl, Chinese carpets, Giotto’s frescoes at Padua, and the masterpieces of Poussin, Piero della Francesca, and Cézanne? Only one answer seems possible—significant form. In each, lines and colours combined in a particular way, certain forms and relations of forms, stir our aesthetic emotions. These relations and combinations of lines and colours, these aesthetically moving forms, I call ‘Significant Form’; and ‘Significant Form’ is the one quality common to all works of visual art. (Bell, 1914)

To construe Bell as being concerned with the rules governing the use of the term ‘art’ or the concept art clearly leads to a reading that is strongly revisionary with respect to Bell’s own conception of the project with which he takes himself to be concerned. Likewise, when Bell claims that art is essentially linked with emotion, leading him to define artworks as objects that cause emotions, Bell is evidently not interested in covertly promoting semantic rules governing the use of the words ‘art’ and ‘emotions’ or the corresponding concepts. Rather, he takes himself as being concerned with a worldly connection which, in his view, holds between artworks and emotions. In this way, Thomasson’s diagnosis, if it were correct, can be seen to generalize well beyond metaphysics, to the social or natural sciences (e.g., economics and biology) as well as to other domains (e.g., art criticism). There is therefore in principle no limit to the scope of the modal normativist’s deflationism.

7 | CONCLUSION

In this paper, we have provided a partial defense of the view we call ‘Socratic essentialism’ (SE), named after its affinity to Plato’s commitment to essentialism as a dialectical presupposition operative in his early Socratic
dialogues. (SE) maintains that some (but not necessarily all) phenomena have essences; that these essences can be rigorously studied and debated by philosophical, scientific, or other means; and that the apparently substantive nature of definitional disputes concerning Socratic ‘What is F?’ questions warrants a commitment to essences. Beyond that, however, (SE) remains neutral and does not require, for example, that essences across the board must be modeled along the lines of Kripke/Putnam-style natural kind essences as intrinsic or microstructural; as hidden or discoverable only by a posteriori means; or as independent of human categorizations or other activities.

Essentialism has been subjected to a wide range of important and serious challenges. Our goal in this paper has been to take up one specific objection, popular among neo-Carnapians, according to which a great majority of apparently substantive definitional disputes are in fact verbal and therefore cannot be used to justify a commitment to ‘heavy-weight’ realist essences. In response to this threat, we argued that the recognition of verbal disputes, as such, by no means undermines our approach; rather, it in fact supplies the Socratic essentialist with more—rather than fewer—essences than initially suspected. By contrast, adopting a ‘light-weight’ construal of the relevant essence-ascriptions as ultimately concerning the rules governing the use of a term or concept, results in an unattractively revisionary conception not only of the subject-matter and methodology of metaphysics and philosophy more broadly (a welcome outcome by neo-Carnapian lights); in addition, the deflationary attitude in question, we argued, will spread beyond its intended target and affect the treatment of apparently substantive definitional disputes in the natural and social sciences (e.g., biology and economics) as well as in other domains (e.g., art criticism).

To develop a more complete picture of the advantages inherent in a commitment to ‘heavy-weight’ realist essences over the neo-Carnapian ‘light-weight’ construal of essence-ascriptions would require us, among other things, to engage more deeply with the epistemology of essences, since the latter point is often mentioned as a major reason to shy away from a realist conception of essences. We therefore take ourselves to have supplied only one, among several, arguments needed for a broader defense of Socratic essentialism. In the end, however, as we hope to demonstrate elsewhere, their explanatory advantages more than outweighs whatever initial qualms we may have had concerning our epistemic access to ‘heavy-weight’ realist essences.

ACKNOWLEDGMENTS
We would like to acknowledge our indebtedness to David Oderberg and Petter Sandstad for their excellent editorial guidance and for putting together this special issue. We received very helpful feedback on earlier versions of this paper from Amie Thomasson, Petter Sandstad, the members of the University of Geneva’s eidos group, and the members of the University of Neuchâtel’s research writing seminar, in particular: Raffaella Antonini, Joshua Babic, Robin Bianchi, Claudio Calosi, Simon-Pierre Chevarie-Cossette, Fabrice Correia, Salim Hirèche, Martin Glazier, Marco Marabello, Kevin Mulligan, and Antoine Taillard. Open access funding provided by Universite de Neuchatel.

ORCID
Kathrin Koslicki https://orcid.org/0000-0002-0868-7199

REFERENCES


**How to cite this article:** Koslicki, K., & Massin, O. (2023). A Socratic essentialist defense of non-verbal definitional disputes. *Ratio*, 00, 1–15. https://doi.org/10.1111/rati.12376