

Essence and definition by abstraction

Bob Hale

Received: 5 November 2017 / Accepted: 13 January 2018 / Published online: 7 March 2018
© The Author(s) 2018. This article is an open access publication

Abstract We may define words or concepts, and we may also, as Aristotle and others have thought, define the things for which words stand and of which concepts are concepts. Definitions of words or concepts may be explicit or implicit, and may seek to report preexisting synonymies, as Quine put it, but they may instead be wholly or partly stipulative. Definition by abstraction, of which Hume’s principle is a much discussed example, seek to define a term-forming operator, such as the number operator, by fixing the truth-conditions of identity-statements featuring terms formed by means of that operator. Such definitions are a species of implicit definition. They are typically at least partly stipulative. Definitions of things, or real definitions, are, by contrast, typically conceived as true or false statements about the nature or essence of their definienda, and so not stipulative. There thus appears to be an obvious and head-on clash between taking Hume’s principle as an implicit and at least partly stipulative definition of the number operator and taking it as a real definition, stating the nature or essence of cardinal numbers. This paper argues that this apparent tension can be resolved, and that resolving it sheds light on part of the epistemology or essence and necessity, showing how some of our knowledge of essence and necessity can be a priori.

Keywords Essence · Definition · Abstraction · A priori knowledge

Professor Bob Hale passed away shortly after completing this paper. He was an invaluable member of our philosophical community and will be sorely missed. This paper is part of a special issue, “New Directions in the Epistemology of Modality” (ed. by A. Mallozzi), which is dedicated to his memory.

1 Introductory remarks: some kinds of definition

There are doubtless different ways in which the expression ‘definition by abstraction’ may be understood.¹ What I mean by it is definition by means of what are now widely known as *abstraction principles*, that is, principles of the form: $\Sigma(\alpha) = \Sigma(\beta) \leftrightarrow Eq(\alpha, \beta)$, where Eq is to be understood as an equivalence relation on entities of the type of α and β and, if all goes well, Σ can be taken to be a function from entities of that type to objects. The principle is understood as equivalent to its universal closure with respect to α and β , which may be first- or higher-order variables. This fixes the order of the abstraction. Thus, to give some well-known examples:

Directions: The direction of line a = the direction of line b iff lines a and b are parallel with the universal closure: $\forall x \forall y (Dir(x) = Dir(y) \leftrightarrow x // y)$ is a first-order abstraction, while:

Hume’s principle: The number of F s = the number of G s iff the F s correspond 1-1 with the G s

with the universal closure: $\forall F \forall G (Nx F(x) = Nx G(x) \leftrightarrow Eq_x(F(x), G(x)))$ is second-order.

At least in the context of the neo-Fregean programme, definition by abstraction is viewed as a form of *implicit definition* which we may employ to establish the meaning of fundamental terms for certain mathematical theories. In the case on which most work has been done, and on which most critical attention has been focused, the idea has been that Hume’s principle may be used to fix the meaning of the number operator $Nx \dots x \dots$, in terms of which the singular term ‘0’, and the predicates ‘...is a cardinal number’, ‘...precedes_ _’, ‘...ancestrally precedes_ _’, and ‘...is a natural number’ may then be defined in well-known ways:

$$\begin{aligned} Cardinal(x) &=_{df} \exists F(x = Nx F(x)) \\ 0 &=_{df} Nx : x \neq x \\ P(x, y) &=_{df} \exists F \exists z (F(z) \wedge x = Nu(F(u) \wedge u \neq z) \wedge y = Nu F(u)) \\ P^*(x, y) &=_{df} \forall F ((F(x) \wedge \forall u \forall v (F(u) \wedge P(u, v) \rightarrow F(v))) \rightarrow F(y)) \\ Natural(x) &=_{df} (x = 0 \vee P^*(0, x)) \end{aligned}$$

Our emphasis has generally been on the idea that in giving the implicit definition, what we do is to *stipulate* that the number operator is to mean whatever it needs to mean—no more, no less—for the sentence, Hume’s principle, to express a truth, when the other expressions involved are taken to have their already established meanings, and the syntax is taken at face value. Much discussion has accordingly focused, quite rightly, upon the questions: Under what conditions, if any, can such a stipulation succeed in fixing a meaning for the definiendum? and: Are those conditions met, in the present case?²

This stipulative conception of how Hume’s principle might be used as a definition is sometimes contrasted with what might be called an *interpretative* or *hermeneutical*

¹ For an illuminating historical discussion, see Mancosu (2016, ch. 1).

² For an extended discussion of such conditions, see Hale and Wright (2000), reprinted as essay 5 in Hale and Wright (2001).

conception, according to which the aim is rather to capture the meaning of an expression as it is already used. This is of a piece with what Quine, in ‘Two dogmas ...’, described as the kind of definition put forward by lexicographers, aimed at reporting pre-existing relations of synonymy, ‘implicit in general or preferred usage prior to [their] own work’ (Quine 1953, p. 24). Whether Hume’s principle could, with any plausibility, be seen as giving a definition of this sort is—in my view—not a question that is easily answered. The answer must be that it cannot, if we take Quine’s description as definitive—i.e. as requiring of an interpretative definition that it present one complete expression as synonymous with another. Such a definition is what would normally be classified as explicit, but Hume’s principle clearly cannot function as a definition in that sense. What Hume’s principle can be used to define, if it can serve as a definition at all, is the number operator; but it presents no expression as *synonymous* with it. If, instead, we take recording a pre-existing synonymy relation as inessential to an interpretative definition, we could consider whether Hume’s principle could be regarded as providing an interpretative definition for the expression ‘the number of ...’ (where the ... holds place for a count noun or noun-phrase), or perhaps of ‘the number of things which ...’ (where the ... holds place for a verb-phrase, i.e. a one-place predicate). Perhaps more plausibly, we could take the question to be whether the definitions of ‘cardinal number’ and ‘natural number’ based on the definition of the number operator by means of Hume’s principle can be regarded as capturing some already established meaning borne by these expressions. My impression is that some have thought the answer to this question is obvious: that they could *not*. I don’t myself think it is so obvious. A case which strikes me as relevantly similar to that of *number* is that of *length*. In that case, we have the noun ‘length’, which may function as a *concrete* noun (as in ‘he was holding a length of rope’), but also as an *abstract* one, as in ‘the length of my yacht is greater than the length of yours’; and we have the cognate expressions ‘the length of...’ and ‘... (is) as long as__’. It seems to me quite plausible that of these expressions, it is the last which is basic or fundamental. Roughly, what we learn first is how to make judgements of the forms ‘x is as long as y’, ‘x is shorter than y’, etc. Only once we have got the hang of these can we, by a kind of abstraction, come to understand talk of lengths—we learn that the length of x is the same as/less than/greater than that of y just when x is as long as/shorter than/longer than y; and by a further kind of abstraction, we come to talk of length and lengths without thinking of them as lengths *of* anything in particular. While there are doubtless differences between talk of length and talk of number, it does not seem to me obviously implausible that similar relations hold between the relational expression ‘...as many as__’, ‘the number of ...’ and ‘number’.

There is, of course, a kind of halfway house between stipulative and interpretative definition—what Carnap called *explication*. As Quine puts it, “In explication the purpose is not merely to paraphrase the definiendum into an outright synonym, but actually to improve upon the definiendum by refining or supplementing its meaning.” (Quine 1953, p. 25). Since both refinement and supplementation require that the pre-explicative meaning of the definiendum be, to some extent, preserved, one might think of explication as a kind of *constrained* or *guided stipulation*. Since the definition of the number operator, and the further definitions based on it, are clearly meant to preserve

certain aspects of ordinary talk and thought about number and counting, there is, it seems to me, good reason to view them as explications in something like Carnap's sense.

In all of these kinds of definition, the aim is to fix the *meaning* of a *word* or other expression. All are thus to be distinguished from, and contrasted with, definition in a more Aristotelian sense, sometimes called *real*, as opposed to *verbal* or *nominal*, definition, where what is defined is *not* the *word* or phrase, but the *thing*—what the word stands for or applies to. The aim of this kind of definition is to capture the *essence* or *nature* of whatever it is that the corresponding word stands for. Such a definition can be assessed as correct or incorrect, not in the sense that it captures or fails to capture the meaning of the word, but according as it is a true or false statement of what it is to be what the word stands for. Some of the theoretical identifications Kripke discusses in *Naming and Necessity* (Kripke 1980), although not expressly presented as such, may plausibly be seen as definitions of this kind:

Water is H₂O

Light is electromagnetic radiation between certain wavelengths

Heat is the motion of molecules

Lightning is an electrical discharge

Obviously these are only rough and ready, but they capture the essential ideas of more accurate scientific statements of the nature of water, light, etc., as contrasted with explanations of the ordinary meanings of the corresponding words.

The central question I wish to discuss concerns the relations between definition by abstraction and definition of this last kind. In particular, what is the relation between Hume's principle, taken as an implicit definition of the number operator, along with the further definitions it supports (of cardinal and natural number predicates, etc.), on the one hand, and the essence or nature of the corresponding entities—the function from concepts or properties to objects, and those objects themselves, i.e. the cardinal numbers in general, and the natural numbers? Can the now familiar sequence of Fregean definitions serve *both* as fixing meanings *and* as definitions in the more Aristotelian sense, as correct statements of the essence of the numbers, etc.?

An obstacle to returning an affirmative answer seems to me to suggest itself right at the outset. Is there not a direct clash between treating something as a definition in accordance with the stipulative conception and treating it as a real definition? Doesn't that require treating the definition, *not* as a *stipulation* of any sort, but as a *true* or *false proposition* aimed at articulating some independently constituted fact about the nature of the worldly correlate of the relevant word?

I believe this apparent tension can be resolved, and that its resolution can shed light both on the notion of a thing's essence or nature—at least as I wish to understand this notion—and on an important part of the *epistemology* of essence. To prepare the ground for the resolution I shall propose, it will be useful to begin with the question: How, in general, are verbal and real definition related?

2 Aristotle and Locke

Although I shall not discuss the views of Aristotle, or those of Locke, at any length, the account I shall present has affinities with both, whilst diverging from both in important respects—so I think it will be useful to comment briefly on those aspects of their views which bear most closely on my main question.

2.1 Aristotle

Aristotle discusses both kinds of definition:

Since a definition is said to be an account of what something is, it is clear that one type will be an account of what its name, or some other name-like account, means – e.g. what triangle means. When we grasp that this exists, we seek why it is. But it is difficult to take anything in this way if we do not know that it exists (Barnes 1993, pp. 93b29–34)

This, as is typical of Aristotle’s logical works, is a very compressed statement. But it is clear enough that here he is talking, in part, about what is sometimes called *nominal* definition. On the plausible interpretation defended by Charles (2000, ch. 2), defining a word is the first preliminary step in a three-stage process which, if successful, culminates in the kind of definition Aristotle sees as the goal of scientific enquiry—an account of the nature or essence of what the word, with that meaning, stands for. The second, intermediate stage consists in verifying that the word, with that meaning, does indeed stand for something—it is only in regard to what exists that we can discover its nature, in Aristotle’s view. As my short quotation illustrates, he seems to identify the final stage—establishing the nature of something—with explaining why it exists. This may at first appear somewhat odd, but we can see roughly what he is driving at by considering some of his examples. Thus a little further on in *Post.An.* B10, he writes:

One definition of definition is the one we have just stated [i.e. saying what the word means]. Another is an account which shows why something exists. Hence the former type means something but does not prove it, whereas the latter will clearly be like a demonstration of what something is, differing in arrangement from a demonstration. For there is a difference between saying why it thunders and what thunder is. In the one case, you will say: Because fire is extinguished in the clouds. But: What is thunder?—A noise of fire being extinguished in the clouds. (Barnes 1993, pp. 93b38–94a7)

What he seems to mean here is that what the word ‘thunder’ means is ‘noise in the clouds’, whereas a definition of what thunder is will incorporate an explanation of this type of noise—it is noise made by fire being extinguished in the clouds. The explanation of thunder is repackaged, as it were, as its definition.

2.2 Locke

Locke discusses essence at length in the *Essay* BkIII, as part of his discussion of words. Central to his discussion is his distinction between ‘nominal’ and ‘real’ essence.

But even before he makes that distinction, the central tenets of his position are laid down:

Words become general by being made the signs of general ideas: and ideas become general by separating from them the circumstances of time and place, and any other ideas that may determine them to be this or that particular existence. By this way of abstraction they are made capable of representing more individuals than one... (Locke 1924, III,3,6)

He that thinks general natures or notions are anything else but such abstract and partial ideas of more complex ones, taken at first from particular experiences, will, I fear, be at a loss where to find them ... This whole mystery of *genera* and *species*, which makes such a noise in the schools, and are, with justice, so little regarded out of them, is nothing else but abstract ideas, more or less comprehensive, with names annexed to them. (op.cit., III,3,9)

... That then which general words signify, is a sort of things; and each of them does that by being a sign of an abstract idea in the mind: to which idea as things existing are found to agree, so they come to be ranked under that name; or, which is all one, be of that sort. Whereby it is evident, that the essence of the sorts ... are nothing else but these abstract ideas. (op.cit. III,3,12)

They [i.e. the abstract ideas, and so the essences of sorts of things] *are the workmanship of the understanding, but have their foundation in the similitude of things.* (op.cit. III,3,13)

After drawing the inevitable conclusion that each distinct abstract idea is a distinct essence, Locke finally comes to the distinction between nominal and real essences:

First, Essence may be taken for the being of anything, whereby it is what it is. And thus the real internal, but generally in substances unknown, constitution of things, whereon their discoverable qualities depend, may be called their essence ...

Secondly, ... [the word 'essence'] has been almost wholly applied to the artificial constitution of *genus* and *species*. It is true, there is ordinarily supposed a real constitution of the sorts of things; and it is past doubt that there must be some real constitution, on which any collection of simple ideas coexisting must depend. But it being evident that things are ranked under names only as they agree in certain abstract ideas to which we have annexed those names, the essence of each genus or sort comes to be nothing but that abstract idea. ... These two sorts of essences, I suppose, may not unfitly be termed, the one the *real*, the other the *nominal*, essence. (op.cit. III,3,15)

Locke mainly discusses essence in connection with general terms for substances: here, Locke claims, real and nominal essence are quite different, and it is nominal essence on which he focuses attention, arguing that it is this which does all the important work, and that talk of real essences is idle and useless—in particular, his view seems to be that it is the abstract idea (i.e. the nominal essence) which defines what it is for something to be of a certain sort, such as *man* or *horse*—we need not trouble ourselves about the generally unknown (and, Locke seems sometimes to suggest,

unknowable) internal constitution of things which causally explains the features we do know about.

2.3 Comments

There are important differences between Aristotle's and Locke's positions, symptomatic of a deep opposition between them. Both focus mainly on 'substances'—in a sense of that term which covers both kinds or sorts of thing, such as horses and meteors, and kinds of stuff, such as gold and water—and on our words for them. But whereas for Aristotle, defining the term or word for something is merely a preliminary stage in the search for essence, for Locke it is the *key*. The important thing for him is the *nominal* essence; he does not deny that there is a real essence, and indeed insists that there must be 'some real constitution, on which any collection of simple ideas coexisting must depend'. But in general, we do not know what that is, and its importance, if known, lies only in its providing a causal explanation of those ideas which collectively constitute the nominal essence. It remains Locke's view that it is possession of those properties corresponding to those simple ideas which make up the abstract general idea—the nominal essence—which determine whether or not something is of a certain kind—and not the 'real internal constitution' on which those properties depend. His view thus stands opposed, not only to Aristotle's, but also to the modern essentialist view promoted by Kripke, according to which the surface properties by which we generally identify natural kinds are not what make them what they are, that being rather the essential properties underlying and explaining presence of the surface properties in normal circumstances. On the Kripkean view, even if it is not naturally possible that things of stuff should be of a certain natural kind—be tigers, or gold, say—whilst lacking certain of those surface features, it is at least metaphysically possible; and conversely, it is metaphysically, even if not naturally, possible that there should be things possessing all the surface features, which are not of that kind (fool's tigers or fool's gold, say). But not on Locke's view. Underlying this difference is a deeper divergence over essential modalities. For Kripke, as for Aristotle, the necessity with which things possess their essential properties is *de re*; but for Locke, it is *de dicto*. To be gold, say, is to conform to a certain abstract general idea which includes having a certain colour and weight, being fusible and fixed (i.e. able to be melted, but not evaporating or losing weight when heated), etc. As we might somewhat anachronistically put it, it is, for Locke, *analytic* that gold is yellow, heavy, fusible, etc. This is why Locke is so adamant that individuals cannot sensibly be taken to have essential properties:

That *essence*, in the ordinary use of the word, relates to sorts, and that it is considered in particular beings no farther than as they are ranked into sorts, appears from hence; that take but away the abstract ideas by which we sort individuals, and rank them under common names, and then the thought of anything essential to any of them instantly vanishes ... Let anyone examine his own thoughts, and he will find, that as soon as he supposes or speaks of essential, the consideration of some species, or the complex idea, signified by some general name, comes into his mind; and it is by reference to that, that this or that quality is said to be essential.

... to talk of specific differences in nature, without reference to general ideas and names, is to talk unintelligibly. For I would ask any one, What is sufficient to make an essential difference in nature, between any two particular beings, without any regard had to some abstract idea, which is looked upon as the essence and standard of a species? All such patterns and standards being laid quite aside, particular beings, considered barely in themselves, will be found to have all their qualities equally essential; and everything in each individual will be essential to it. Or, which is more, nothing at all. (op.cit. III,6,4–5)

There could not be a clearer rejection of *de re* necessity. Locke's position is, in effect, an anticipation of that championed by Quine some two and a half centuries later: necessity resides, not in the things we talk about, but in the way we talk about them.³

Although, in my view, Locke's central claims about essence—driven by his animus against scholastic claims about it, cf. 'This whole mystery of *genera* and *species*, which makes such a noise in the schools'—are mistaken, what he says about nominal and real essence in relation to *non-substance* terms merits separate comment. Immediately following the section in which he introduces his distinction, he remarks:

Essences being thus distinguished into nominal and real, we may farther observe, that in the species of simple ideas and modes, they are always the same; but in substances, always quite different. Thus a figure including a space between three lines, is the real as well as the nominal essence of a triangle; it being not only the abstract idea to which the general name is annexed, but the very *essentia*, or being of the thing itself, that foundation from which all its properties flow, and to which they are all inseparably annexed. (op.cit. III,3,18)⁴

The kind of coincidence of nominal with real essence to which Locke here draws attention—or, what comes to essentially the same thing, agreement between *verbal*

³ Mackie (1974) pointed out that Locke certainly recognises the possibility of adopting a view like Kripke's, according to which our intention in using a term such as 'gold' is to refer to the substance whose internal constitution underlies and explains the surface properties which make up our abstract general idea of gold (its nominal essence), and allows that people may use the word with that intention (cf. "...the mind...makes them [words for substances], by a secret supposition, to stand for a thing, having that real essence", even when we do not know what the real essence is; also "there is scarce any body in the use of these words, but often supposes each of those names to stand for a thing having the real essence, on which these properties depend", Locke 1924, III,10,§18); but he makes it clear that he thinks this practice is misguided and misleading.

⁴ In Selby-Bigge's abridged edition of Locke's *Essay*, the continuation might seem to indicate that Locke is less firmly opposed to essentialism as it is usually understood than I have claimed. It runs: "But it is far otherwise concerning that parcel of matter which makes the ring on my finger, wherein the two essences are apparently different. For it is the real constitution of its insensible parts, on which depend all those properties of colour, weight, fusibility, fixedness, etc., which makes it to be gold, or gives it a right to that name, which is therefore its nominal essence." The subject-main verb agreement virtually obliges us to read this sentence as asserting that it is the real constitution [rather than the properties of colour, etc.] which makes something to be gold. In fact, this is a catastrophic editorial blunder. Locke's original text reads: "For it is the real constitution of its insensible parts, on which depend all those properties of colour, weight, fusibility, fixedness, etc., which are to be found in it, which constitution we know not, and so having no particular idea of, have no name that is the sign of it. But yet it is its colour, weight, fusibility, fixedness, etc., which makes it to be gold, or gives it a right to that name, which is therefore its nominal essence." (see, e.g., Locke 1924, vol. 1, p. 451). Selby-Bigge's compression completely changes Locke's clear meaning.

and *real* definition—seems to me quite certainly to occur, and indeed to be well illustrated by Locke’s own example. To give another, if we ask what it is to be a circle, a good answer is that it is a collection of all and only those points in a plane equidistant from some fixed point in the plane; but this answer is every bit as good, as an answer to the question: What does the term ‘circle’ mean? Similarly, to be a vixen is to be an adult female fox, and that is what ‘vixen’ means. Examples are easily multiplied. This happy coincidence, where real and verbal definitions may be given by the same words, is the starting point for my own answer to my leading questions.⁵

3 A priori knowledge of essence

In many cases, then—or so I claim—a correct definition of what a word stands for may serve as a good definition of the word itself. But in many other cases, it will not. For example, a correct definition of the stuff, water, is probably something along the lines of: the liquid form of a substance composed of molecules consisting of two hydrogen atoms bonded with one oxygen atom; but this is not a good definition of the word ‘water’. You will easily think of other examples. There may also be—indeed, it seems to me that there are—examples where there may be no clear answer whether a correct definition of a thing is also a correct definition of the word, in part because word-meanings change over time, so that what starts out being a discovery about something’s nature may become part of what a word for it means. This might be the case with tigers, for example. A tiger is a large carnivorous mammal of the cat family, of the genus *panthera*, distinguished from other species by certain further characteristics, including its striped skin and coat. I doubt that the *whole* zoological definition would, even now, amount to a correct definition of the word, as used by most speakers. But it is quite plausible that it is now part of the word’s meaning that tigers are mammals and that they are large cats, even if it wasn’t when the word was first used.

The coincidence of correct verbal and real definition helps to explain how it is that, in some cases, we may have a priori knowledge of something’s essence or nature. The

⁵ Locke’s general claim, that nominal and real essence are the same for all simple ideas and modes cannot, it seems to me, be entirely right. To give an admittedly debatable example, the idea of redness would rank for him as a simple idea, but being red (as a property of surfaces) might plausibly be taken to consist in the capacity to reflect light in the range 640–720 nm. To give another, perhaps less debatable, example—one which Locke himself explicitly discusses—the idea of light is simple. But while he would have to say in this case nominal and real essence are the same, it is plausible that, to the contrary, light may be defined, even if the word for it, and the idea that is its meaning (for Locke), cannot be. Locke chides “those who tell us, that light is a great number of little globules, striking briskly on the bottom of the eye”—he grants that they “speak more intelligibly than the schools”, but complains:

but yet these words ever so well understood would make the idea the word light stands for no more known to a man that understands it not before, than if one should tell him, that light was nothing but a company of little tennis-balls, which fairies all day long struck with rackets against some men’s foreheads, whilst they passed by others. (op.cit. III,4,10)

Locke is of course right that the particle theory of light doesn’t explain what the word ‘light’ means, as used by ordinary speakers of English. But that, contrary to what Locke is forced to say by his identification of its nominal with its real essence, does not mean that it cannot be a good definition of what light is.

kind of explanation I have in mind is most readily illustrated by reference to cases in which we can give an *explicit* definition of a word. Thus we can define the noun ‘square’ by:

A square is a plane figure bounded by four straight sides of equal length meeting at right-angles

Such an explicit definition can always be rewritten as a biconditional:

$\forall x(x \text{ is a square} \leftrightarrow x \text{ is a plane figure bounded by four straight sides of equal length meeting at right-angles})$

The definition tells us that being a plane figure, etc., is a necessary and sufficient condition for the noun ‘square’ to be correctly applied to something. But if the noun ‘square’ is correctly applicable to something, that thing is a square. So anyone who is in a position to give this definition of the word can recognize that being a plane figure, etc., is a necessary and sufficient condition for being a square—that is, that that is *all there is* to being a square. But to know that much is to know what it is to be a square. On the reasonable assumption that if all one needs to know, if one is to know a priori that p , is the meaning of some sentence S which expresses the proposition that p , such knowledge is a priori.⁶ What could be easier?

Of course, we cannot always give explicit definitions. But I can see no reason why this simple and straightforward model for a priori acquisition of knowledge of essence should not be extended to other cases—both those in which, although we cannot explicitly define a word, we can give an *implicit* definition, and others in which neither kind of definition can be given. These cases can be illustrated by reference to our understanding of logical constants. We cannot define the truth-functional connectives explicitly. But provided that the meanings of some of them can be taken as already understood, we can implicitly or contextually define others. Thus taking words for negation and disjunction as already understood, we can implicitly define the truth-functional conditional by stipulating that $A \rightarrow B$ is to be true iff $\neg A \vee B$ is. \rightarrow stands for a certain function from propositions to propositions. Someone who understands the implicit definition is able to recognize that the function for which it stands is that function which takes an ordered pair of propositions to a true proposition iff the first of them is false or the second is true. To know this is to know the essence of the truth-functional conditional. Supposing negation and disjunction to be our basic connectives, we cannot define either of them, either explicitly or implicitly. But we can learn what they mean. Anyone who knows what truth-functional disjunction means knows that $A \vee B$ is true iff at least one of the pair A, B is true. This knowledge of what the word means equips them to recognize that truth-functional disjunction is that function from propositions to propositions which takes a pair of propositions to a true proposition iff at least one of the pair is true.

If I am right, there are many cases in which a priori knowledge of essence can be based upon knowledge of the meanings of relevant words. Definition by abstraction can be seen as a special case. Hume’s principle, construed as an implicitly definitional

⁶ Of course, knowledge of the meanings of S ’s ingredient expressions and its semantically significant structure will be a posteriori. But any sensible account of a priori knowledge will allow that.

stipulation, serves—if all goes well—to fix the meaning of the number operator. Once that is done, we may explicitly define

$$x \text{ is a cardinal number} \leftrightarrow_{\text{def}} \exists F(x = Nx Fx)$$

and

$$0 =_{\text{def}} Nx(x \neq x)$$

and finally

$$x \text{ is a natural number} =_{\text{def}} x = 0 \vee P^*(0, x)$$

where P^* is ancestral predecession, where, as Frege himself proposed, predecession is defined by:

$$Pxy =_{\text{def}} \exists F \exists z (y = NuFu \wedge x = Nu(Fu \wedge u \neq z))$$

and its ancestral by:

$$P^*xy = \forall F ((Fx \wedge \forall u \forall v ((Fu \wedge Puv) \rightarrow Fv)) \rightarrow Fy)$$

On the basis of these definitions one can gain knowledge a priori of the nature of cardinal and natural numbers.

4 Definition and essence

4.1 An obvious lacuna

If I am right, knowledge of essence in the kind of case in which my proposed explanation applies is much more easily obtained than in cases in which only a posteriori knowledge is possible. No painstaking empirical investigation or scientific breakthrough is required. Essence is, as it were, open to view. I think this explanation of how we may get a priori knowledge of essence is good as far as it goes. But it does not go far enough. There is, in particular, at least one further question which clamours for an answer: The explanation tells us how, when the definition of a word and a real definition of the thing for which it stands coincide, we can gain knowledge of the nature of the thing from our grasp of the meaning of the word. But how is this happy coincidence—between the definition of the word and the essence of the thing—itsself to be explained? The meanings of words are, in some suitably broad sense, a matter of convention. We fix them, either expressly by explicit or implicit definition, or in less formal ways, by our ongoing linguistic practice. So how does it come about that in cases of the kind I have been discussing, we fix them so that they match up so well with the essence or nature of the things the words stand for?

4.2 Direction of fit

Put like that, the question is liable to appear intractable. For putting it like that assimilates the cases with which we are concerned to those involving natural kinds, which, we assume, have independently constituted natures which we seek to capture by investiga-

tion a posteriori. If we think of our cases on this model, preserving the same direction of fit, it may well seem that achieving fit, without empirical investigation, requires some mysterious kind of pre-established harmony. But that is, fairly obviously, the wrong picture. *To a first approximation*, we might say that the right picture *reverses* the direction of fit: it is not that we somehow miraculously fix meaning so that it conforms to an independently pre-determinate essence—instead, in these cases, *we fix essence by fixing meaning*. I will explain why this can only be at best a first approximation shortly.

If this alternative picture is approximately right, it immediately resolves the apparent tension noted previously, between endorsing the stipulative conception of how Hume's principle may be used to define the number operator and regarding it and the further definitions based upon it as revealing essence. The tension results from a false assimilation to cases in which we seek to capture the essence of natural kinds. Cardinal and natural numbers are not *natural* kinds—at least not in the sense in which gold and tigers are. They are, as one might say, *artificial* kinds.

4.3 Some objections answered

That description risks misunderstanding of a sort that relates closely to a likely cause of resistance to the proposal that we fix essence by fixing meaning. If things of a certain sort—numbers, say—are an artificial kind, doesn't that mean that they are mere artefacts, which would not exist, were it not for our activity? Mustn't the view that we fix essence by fixing meaning carry with it an intolerable sacrifice of objectivity?

Soberly understood, the claim that, in certain cases, we fix essence need have no such unpalatable consequences. To explain why not, I shall draw on two points—one is a point about the specific character of definition by abstraction; the other is a more general point about properties.

The specific point about abstraction is one which Hale and Wright have emphasized on many occasions. Definition by abstraction *creates no objects*. What, if successful, it creates is only a *concept*. Or to put it more accurately, what it does is to fix the meaning of a term-forming operator, such as the number operator. It does this by fixing the condition for identities configuring singular terms formed by applying that operator to suitable arguments to be true. Such identities are to have the same truth-conditions as those of corresponding statements asserting those arguments to stand in a certain equivalence relation. The existence of referents for those terms—numbers, say—is required by the truth of such identities. If the definition succeeds, such identities are true iff the corresponding statements involving the relevant equivalence relation are. It follows that so long as it is a matter of objective, mind-independent fact that, say, there is a one-one correlation between the *F*s and the *G*s, the existence of the number of *F*s and the number of *G*s (and their identity) is an equally objective, mind-independent fact.

That point may be thought enough, by itself, to dispel any reasonable concern about loss of objectivity. But my more general point about properties is worth making, because it illuminates the kind of existence numbers, and other pure abstract objects, enjoy. As just noted, if Hume's principle can be used to define the number operator,

standing for a certain function from properties to objects, the existence of numbers depends only upon there being properties which are one-one correlated. No more is required for the existence of numbers than that of first-level properties and a certain second-level equivalence relation. According to the *abundant or deflationary conception* of properties and relations, no more is required for the existence of any purely general property or relation than that there *could* be a predicate with the appropriate satisfaction condition.⁷ Provided that this very weak existence condition is met for some suitable purely general properties and the purely general relation of one-one correlation, no more is required for the existence of cardinal numbers. This brings out the sense in which numbers are ‘metaphysically lightweight’ objects. Numbers no more depend upon us for their existence than does the relation of one-one correlation or the purely general properties which stand in that relation. They all depend for their existence on the possibility of expressions of the appropriate logico-syntactic type having them as their semantic values, but the requisite facts about possibility are themselves in no way dependent upon us or any other thinkers and speakers. Provided that it is agreed that the kind of possibility involved in formulating the existence condition for general properties conforms to S5 principles, it can easily be shown that their existence, and hence that of numbers, is absolutely necessary.⁸ To put the point slightly paradoxically, whilst we fix *what it is to be* a cardinal or natural number, *their being* is absolutely necessary and hence quite independent of merely contingent beings, such as—on one widely but, of course, not universally accepted view—we are!

⁷ See Hale and Wright (2009, pp. 207–218), and for a somewhat fuller discussion, Hale (2015, ch.1, esp. 1.12).

⁸ Let ϕ be any purely general property, and let p be the statement that ϕ exists, and q the statement that there exists a predicate standing for ϕ . Then according to the deflationary account:

$$p \leftrightarrow \Diamond q$$

Further, this statement, if true, will be so as a matter of necessity, at least on the deflationary account, so that:

$$(a) \quad \Box(p \leftrightarrow \Diamond q)$$

But (a) entails both:

$$(b) \quad p \leftrightarrow \Diamond q$$

and

$$(c) \quad \Box p \leftrightarrow \Box \Diamond q$$

and, by S5, (c) entails

$$(d) \quad \Box p \leftrightarrow \Diamond q$$

But from (b) and (d), it follows by the transitivity of the biconditional that

$$(e) \quad p \leftrightarrow \Box p$$

Since ϕ may be any purely general property, it follows that each purely general property there is exists as a matter of necessity.

4.4 Direction of fit again

Against the backdrop of the abundant conception of properties (and objects), it is easier to pinpoint what is wrong, and also what is right, in the suggestion that we fix essence by fixing meaning.⁹ What is wrong with it is that it does insufficient justice to the fact that the properties that constitute a thing's essence or nature exist quite *independently* of any definitional activity on our part. We *fix the meaning* of the word—'square' say—by stipulating that to be square, a plane figure must be composed of four sides of equal length meeting at right angles. But that *there exists* such a *property*—the property of being a figure so composed—is entirely *independent* of any such definition being given (though not, on the abundant conception, independent of the *possibility* of such a definition). What is right about it—or at least very nearly right—lies in the suggestion that there is little or no room for error. But the explanation how we achieve such a happy match between our definition and the essence is not that we fix the essence to suit our definition; it is rather that, *given the abundant conception, we cannot very well miss the target*. For according to that conception, provided that there could be a predicate with the appropriate satisfaction-condition, the existence of the corresponding property is guaranteed. Thus provided only that our purported definition is in good order, it cannot fail to pick out the requisite defining property.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix: Note on Dummett on mathematical objects

In his paper 'What is mathematics about?' (Dummett 1993, pp. 429–445), Michel Dummett writes:

In order to confer sense upon a general term applying to concrete objects – the term “star” for example ... we consider it enough that we have a sharp criterion for whether it applies to a given object, and a sharp criterion for what is to count as one such object – one star, say – and what as two distinct ones: a criterion of application and a criterion of identity. The same indeed holds true for a term, like ‘prime number’, applying to mathematical objects, but regarded as defined over an already given domain. It is otherwise, however, for such a mathematical term as “natural number” or “real number” which determines a domain of quantification. For a term of this sort, we make a further demand: namely, that we should ‘grasp’ the domain, that is, the totality of objects to which the term applies, in the sense of being able to circumscribe it by saying what

⁹ I am much indebted here to Joachim Horvath, both for pressing the objection originally, and for subsequent discussion of it, in which it emerged that the line I propose here has close affinities with one taken in his dissertation. See Horvath (2011).

objects, in general, it comprises – what natural numbers, or what real numbers, there are.

The reason for this difference is evident. For any kind of concrete object, or abstract object whose existence depends upon concrete objects, external reality will determine what objects of that kind there are; but what mathematical objects there are within a fundamental domain of quantification is supposed to be independent of how things happen to be in the world, and so, if it is to be determinate, *we* must determine it. (p. 438)

This passage is apt to seem somewhat puzzling. We may well agree with Dummett that the composition of a fundamental mathematical domain, such as the natural or the real numbers, should be—as he puts it—*independent of how things happen to be in the world*, but wonder what exactly is meant by the suggestion that if it is to be determinate, *we* must determine it, and be reluctant to agree to it.

For Dummett, the remarks I have quoted are a prelude to arguing that fundamental mathematical domains—including those of the real and even the natural numbers—are indefinitely extensible, and, on the back of that, that we may only legitimately employ constructive, non-classical principles in reasoning about them. This is hardly the place to take on those arguments. Instead, I shall try to sketch a way in which the ideas of the preceding section might be deployed to preserve a modest form of realism about the composition of such domains. I shall, however, need to take issue with one of Dummett's early claims in making his case. He writes:

One the face of it, indeed, a criterion of application and a criterion of identity do not suffice to confer determinate truth-conditions on generalizations involving some general term, even when it is a term covering concrete objects: they can only give them content construed as embodying a *claim*. So understood, an existential statement amounts to a claim to be able to give an instance; a universal statement is of the form “Any object to which the term is recognized as being applicable will be found to satisfy such-and-such a further condition”. (p. 438)

Since Dummett offers nothing in support or even explanation, I can only conjecture that he is taking criteria of application and identity in an *epistemological* sense, so that they embody the means whereby judgements of application and identity may be made, in contrast with conditions for the *truth* of such judgements. But, at least in the case of the natural numbers, I can see no compelling reason why we should not take our Fregean definitions, based ultimately upon an implicit definition of the number operator by means of Hume's principle, as giving necessary and sufficient conditions for the application of the predicate ‘ ξ is a natural number’ and for the identity or distinctness of its instances, and so (together, of course, with any other definitions of arithmetical terms which may be needed) providing *truth-conditions* for number-theoretic statements.¹⁰

¹⁰ To be sure, Dummett believes the impredicativity of Hume's principle—unavoidable if it is to serve as a foundation for arithmetic—means that it cannot be acceptable. But Wright and I have endeavoured to deal with this objection elsewhere. See Hale (1994, §6), and Wright (1998), both reprinted in Hale and Wright (2001).

If we may so take them, we may agree with Dummett that *we* determine what lies within the domain of natural numbers, *in the sense* that we fix the *concept* of *natural number*, and thereby fix *what it is to be* a natural number, whilst holding that what objects that domain comprises is not itself of our making in any sense that conflicts with a modest realism, but is determined by how things stand in the world, independently of us and our activities. To be sure, it is entirely independent of *contingent* features of the world, just as Dummett observes it should be. For the existence of the properties and the relation of one-one correspondence, and that of the operation whose values the natural numbers are for properties bearing that relation to one another, involved in the Fregean definitions, is—on the abundant or deflationary theory—a matter of *necessity*, entirely independent of us.

It is a further—and in my view, much more difficult—question, whether the same may be said of the domain of real numbers. Of this case, Dummett writes:

Cantor's celebrated diagonal argument to show that the set of real numbers is not denumerable has precisely the form of a principle of extension for an indefinitely extensible concept: given any denumerable totality of real numbers, we can define, *in terms of that totality*, a real number that does not belong to it. (p. 442)

To this, there is an obvious response (as Dummett himself notes): to show that no *denumerable* totality of real numbers contains them all is not to show that there is no determinate or definite totality of real numbers, for there may be, as Dummett's realist opponents suppose, a *non-denumerable* one. Simply to assume that any definite totality must be denumerable is to beg the question against them. Dummett seeks to forestall this objection:

The argument does not show that the real numbers form a non-denumerable totality unless we assume at the outset that they form a determinate totality comprising all that we shall ever recognize as a real number: the alternative is to regard the concept *real number* as an indefinitely extensible one. (ibid)

But it is unclear, at best, that the realist must *prove* that the real numbers form a definite totality. On the face of it, she may protest, there is a set—definite totality—comprising all and only the finite cardinals, i.e. the natural numbers—it comprises exactly 0 together with those cardinal numbers which bear the ancestral of the immediate successor relation to it. Similarly, the set of reals has as its members precisely those numbers which are determined by Dedekind cuts in the rationals (or by Cauchy sequences of rationals), so that here too there is a *prima facie* case for taking them to form a definite totality. Of course, there is a parallel *prima facie* case for supposing there to be a set of all ordinals, but as is shown by Burali-Forti's paradox, there can be no such definite totality, *denumerable or otherwise*. However, there is no parallel argument to show that the reals cannot form a definite totality.

The issue is, as Dummett sees, one of where the burden of proof lies. He grants that such issues are always difficult to resolve, but thinks that in the present case it is clearly the realist who has begged the question and must shoulder the burden—her constructivist opponent 'does not *assume* ... that any totality of which it is possible to form a definite conception is at most denumerable; he merely has as yet no conception

of any totality of higher cardinality’ (p. 443). But it is *not* clear. We can state a necessary and sufficient condition for something to be a real number. It must be granted that this does not *guarantee* the existence of a totality comprising all and only the real numbers—so much is the moral of the Burali-Forti. ‘... the characterization of an indefinitely extensible concept’, Dummett claims, ‘demands much less than the once-for-all characterization of a determinate totality’ (p. 442). But he gives no hint as to what more may be required, and it is unclear what further positive requirement, beyond provision of necessary and sufficient conditions for membership, might be imposed—as distinct from the absence of any defeating argument, paralleling Burali-Forti’s paradox, to show that the assumption of a definite totality leads to contradiction.

References

- Barnes, J. (Ed.). (1993). *Aristotle posterior analytics* (2nd ed.). Oxford: Clarendon Press.
- Charles, D. (2000). *Aristotle on meaning and essence*. Oxford: Clarendon Press.
- Dummett, M. (1993). *The seas of language*. Oxford: Clarendon Press.
- Hale, B. (1994). Dummett’s critique of Wright’s attempt to resuscitate Frege. *Philosophia Mathematica*, 2, 122–47.
- Hale, B. (2015). *Necessary beings: An essay on ontology, modality, and the relations between them* (2 paperback ed.). Oxford: Clarendon Press.
- Hale, B., & Wright, C. (2000). Implicit definition and the a priori. In P. Boghossian & C. Peacocke (Eds.), *New essays on the a priori* (pp. 286–319). Oxford: Clarendon Press.
- Hale, B., & Wright, C. (2001). *The reason’s proper study: Essays towards a Neo-Fregean philosophy of mathematics*. Oxford: Clarendon Press.
- Hale, B., & Wright, C. (2009). The metaontology of abstraction. In D. Chalmers, et al. (Eds.), *Metametaphysics* (pp. 178–212). Oxford: Clarendon Press.
- Horvath, J. (2011). *In defence of conceptual analysis*. Ph.D. thesis, University of Cologne.
- Kripke, S. (1980). *Naming and necessity*. Oxford: Basil Blackwell.
- Locke, J. (1924). *Essay concerning human understanding*. Oxford: Clarendon Press.
- Mackie, J. L. (1974). Locke’s anticipation of Kripke. *Analysis*, 34, 177–80.
- Mancosu, P. (2016). *Abstraction and infinity*. Oxford: Oxford University Press.
- Quine, W. V. (1953). Two dogmas of empiricism. In W. V. Quine (Ed.), *From a logical point of view*. Cambridge: Harvard University Press.
- Wright, C. (1998). On the (harmless) impredicativity of Hume’s principle. In M. Schirn (Ed.), *Philosophy of mathematics today* (pp. 339–68). Oxford: Clarendon Press.