Category theory as an autonomous foundation

Abstract

Does category theory provide a foundation for mathematics that is autonomous with respect to the orthodox foundation in a set theory such as ZFC? We distinguish three types of autonomy: logical, conceptual, and justificatory. Focusing on a categorical theory of sets, we argue that a strong case can be made for its logical and conceptual autonomy. Its justificatory autonomy turns on whether the objects of a foundation for mathematics should be specified only up to isomorphism, as is customary in other branches of contemporary mathematics. If such a specification suffices, then a category-theoretical approach will be highly appropriate. But if sets have a richer 'nature' than is preserved under isomorphism, then such an approach will be inadequate.

A number of philosophers of mathematics have recently debated the claim that category theory provides a foundation for mathematics that is autonomous with respect to the orthodox foundation in set theory ([26], [10], [29], [4], [14], [33], [1]). The debate has yielded progress: after some initial confusion, the particular theories from within category theory that are proposed as foundations have been identified precisely, and in some cases the autonomy of these theories with respect to the orthodox foundation has been defended—at least for one sort of autonomy. However, there are other sorts of autonomy that have not been considered in much detail. We wish to introduce a distinction between three types of autonomy, which we call *logical autonomy*, *conceptual autonomy*, and *justificatory autonomy*. The debate so far has been concerned almost exclusively with the first sort of autonomy. Yet all three are required before a foundation can claim genuine independence from the set-theoretic orthodoxy.

In the second half of this paper, we focus on one of the putative category-theoretic foundations that has been proposed and we make a case study of it, arguing that considerations that arise with respect to this foundation will arise with respect to the numerous others. We argue that this foundation can claim logical autonomy with respect to orthodox set theory.

We then explore the possible arguments that could be made for or against its conceptual and justificatory autonomy. We argue that the debate turns crucially on whether the objects of a foundation for mathematics can or indeed should be specified only up to isomorphism, as is customary in other branches of contemporary mathematics. In particular, if sets should be characterized only up to isomorphism, then a category-theoretical approach will be highly appropriate; whereas if sets have a richer 'nature' than is preserved under isomorphism, then such an approach will be inadequate.

It is often said that category theory provides a unificatory language in which all of mathematics may be stated and in which the important connections between key concepts in different disciplines is most perspicuously revealed. We will have nothing to say about this claim, except to emphasise that it is independent of the questions we address here.

1 An overview of the debate

Many category theorists, including Saunders Mac Lane, William Lawvere, Steve Awodey, and John Bell, have claimed that category theory or topos theory has the resources to provide a foundation for all of mathematics (in a sense yet to be specified) that is independent of the orthodox foundation in a set theory such as ZFC ([20], [26], [1], [5]). Call the proponent of such a view a *categorist*. Against the categorist, Solomon Feferman and Geoffrey Hellman have raised two main objections: the Mismatch Objection and the Logical Dependence Objection ([10], [14]).

The Mismatch Objection maintains that neither category theory nor topos theory are the right sort of thing to act as a foundation. After all, a foundational theory must make assertions, and in particular existential assertions. A foundational theory should provide us with a theory of the objects of mathematics; and it must consist of assertions of the existence of those objects, as well as an account of the relations in which they stand. In Hellman's terminology, its axioms must be *assertory* ([15]). However, neither the Eilenberg-Mac Lane axioms for category theory nor the Lawvere-Tierney axioms for topos theory have this form. Rather, when taken together, each set of axioms provides a *definition* of a certain sort of mathematical structure: in the one case, a mathematical structure called a *category*; in the other case, a *topos*. In this sense, they are akin to the axioms of the theory of groups (or the theory of rings or indeed any other algebraic theory): the group axioms together provide a definition of what it is for a mathematical structure to be a group; they do not make assertions about the existence of groups. In Hellman's terminology, each of these sets of axioms—the topos axioms, the category axioms, and the group axioms—is *algebraic-structural* ([15]). But no such collection of definitions can form the foundation for a discipline. At most, a definition can take a class of objects that are held to have foundational significance, and delimit from amongst them a particular subclass whose members may be thought to provide the subject matter of mathematics. But this will not suffice for a foundation for mathematics. What is required for a foundation is an assertory theory of the objects of foundational significance. This is not supplied by the axioms for a category, a topos, or a group. Thus, according to Feferman and Hellman, there is a mismatch between the foundational role that the categorist would like her theory to play, and the sort of theory that she claims plays it.[1]

So the Mismatch Objection identifies a problem with the form of the proposed foundation for mathematics in category theory. By contrast, the Logical Dependence Objection attacks the relationship between that proposal and the orthodox foundation for mathematics in set theory. It claims that category theory and topos theory are not logically autonomous with respect to set theory. Rather, they depend logically in two different ways upon a prior theory of sets and functions, which thus provides the true foundation for mathematics.

Firstly, as we saw above, the axioms for category theory or topos theory provide definitions. These definitions are given in the form of necessary and sufficient conditions for two *classes* (the class of objects and the class of arrows) together with three *functions* (the domain, codomain, and composition functions) to count as a category or topos. According to the Logical Dependence Objection, it follows that the theory of categories is simply a theory of classes and functions—two classes and three functions to be precise—that satisfy certain conditions. The theory of classes and functions is thus logically prior to the theory of categories. In Feferman's helpful analogy, category theory stands to the theory of classes as the definition of linear transformation stands to the definition of vector spaces: in both cases, it is not possible to state the former without having previously stated the latter (152-3, [10]).

Secondly, whilst it is correct in the Mismatch Objection to claim that the axioms of

[1]The deductivist approach to mathematics would deny that there is any mismatch. This approach maintains that mathematics consists of a collection of conditionals whose antecedents are conjunctions of definitional axioms, and whose consequents are theorems concerning the sort of mathematical structures thereby defined. However, the deductivists who propose category theory as the correct framework in which to state these conditionals are vulnerable to the same objections as those who prefer the set-theoretic framework. We do not consider their position here, but see [2] and [15] for both sides of the debate.

category theory and topos theory make no existential assertions, textbook presentations of these subjects do.[2] These existential assertions concern particular categories or toposes, such as Grp, the category whose objects are set-sized groups and whose arrows are the group homomorphisms between them. But the objects of these particular categories are standard mathematical structures, each given as a set together with various functions and relations on that set. So again category theory and topos theory depend logically on a prior theory of classes and functions in order to ground their existential assertions. They depend on a theory of classes officially, since their axioms serve to define when certain collections of classes and functions are to count as together forming a category or a topos; and they depend on such a theory unofficially to provide witnesses for the existential claims made in their textbook presentations.

McLarty has responded to both objections on behalf of the categorist. He agrees that these objections would refute anyone who tried to provide a foundation for mathematics in the *general* theory of categories or the *general* theory of toposes [33]. But he denies that anyone has ever proposed such a foundation. Responding specifically to Hellman's claim that category theory and topos theory make no assertions, he replies:

> This is quite true of the category axioms *per se*, and of the general topos axioms. That is why no one offers them as foundations for mathematics. (45, [33])

Rather, he claims, the foundational theories proposed by the categorists are all theories of some *particular* category or topos. More precisely, according to McLarty, three such theories have been proposed as foundations for mathematics. They are these: SDG, Lawvere and Kock's theory of the category of smooth spaces and the smooth maps between them ([21], [18]);[3] CCAF, Lawvere's theory of the category of categories and the functors between them ([20], [31]); and ETCS, Lawvere's theory of the category of abstract sets and arbitrary mappings between them ([19], [23]). Furthermore, there are putative foundational theories proposed by categorists over and above those listed by McLarty. For instance: Makkai's

[2] Of course, the topos axioms do entail *hypothetical* existential assertions: for instance, they entail that, *if* a given category is a topos, *then* there exists a terminal object that belongs to it. But these are not the sort of existential assertions that a foundation must contain. Similarly, the group axioms entail hypothetical existential assertions: for instance, they entail that, *if* a given structure is a group, *then* there is an identity element that belongs to it. But again, this is not the sort of existential assertion that a foundation must contain.

[3] Though this theory has only ever been proposed as a foundation for a particular part of mathematics, namely, differential geometry.

SFAM [28], Bénabou's theory of fibered categories [7], and Bell's theory of local set theories [6] all claim to provide a foundation for all of mathematics, just as ETCS and CCAF do; while Hyland's theory of the effective topos [17] and the theory of a particular Abelian category [11] are amongst those that attempt to provide a foundation for a particular part of mathematics, just as SDG does—in these cases, computability theory and homological algebra, respectively. Furthermore, there is a treatment of set theory given in category-theoretic terms that is quite different from ETCS. It is called *algebraic set theory*, and it is the algebraic study of models of ZF set theory and various of its interesting subsystems [3]. We do not consider it here, since it does not provide a significantly different foundation from that given by ZF set theory itself. Again, this is not to disparage its mathematical use and power. It is simply to note that its subject matter is close to ZF—universes of sets structured by the membership relation. Thus, it is not a rival foundation. With these theories in hand, let us consider how they fare with respect to the Mismatch and Logical Dependence Objections.

The Mismatch Objection says that a particular proposed foundation is not the right sort of theory to provide a foundation because it is not assertory. And, in the cases considered by Feferman and by Hellman, the reason these theories are not assertory is that they do no more than provide definitions of technical terms, such as 'category' or 'topos'. But of course there are other ways in which a theory can fail to make assertions, and thus fall foul of the Mismatch Objection. For instance, it might be that the theory is not contentful. This would be the problem, for instance, with the theory that consists of the following sentence: 'The mome raths outgrabe'. The reason that this theory lacks content is that there is no way of identifying, independently of the theory, what mome raths are, or what it is to outgrabe. Without identifying the subject matter of the theory in such a way, the statements of the theory do not make contentful assertions. A similar worry arises for some of the theories presented above. Take, for instance, the theory of the effective topos. No identification of the subject matter of this theory has been given that is both independent of the theory itself and also independent of set theory. Of course, it is possible to define the theory's subject matter as comprising a certain sort of set. But then the theory will answer the Mismatch Objection in a way that makes it clearly susceptible to the Logical Dependence Objection. If the subject matter of the theory is simply a certain kind of set, then it is not possible to state the theory fully without reference to sets. So, in this case, we face a dilemma. The theory is either vulnerable to the Mismatch Objection, or it avoids that objection in a way that makes it vulnerable to the Logical Dependence Objection. We emphasise again that this is not to impugn the mathematical value of Hyland's definition; it is only to question its ability to provide an autonomous foundation.

The same problem arises for the theory of a particular Abelian category. And it also seems to arise for Lawvere's CCAF, which is the theory of categories themselves. CCAF asserts the existence of certain categories and describes some of the functors between them. The theory is not definitional, so it is not vulnerable to the Feferman-Hellman version of the Mismatch Objection. But if it is to make a contentful assertion, we need to be able to identify its subject matter—namely, categories—independently of the theory. Of course, we can do this by taking a category to consist of a set of objects and a set of arrows and three functions with particular properties, as is done in the original definition, due to Eilenberg and Mac Lane. But then CCAF becomes vulnerable to the Logical Dependence Objection.[4] Is there an alternative account of categories that avoids this problem? Certainly, there are a number of different definitions of the notion. For instance, in any category with sufficient structure, we can define the notion of an 'internal category', which is intended to generalise the original set-theoretic definition. But of course, this account, while mathematically fruitful, is question-begging as an attempt to identify the subject matter of CCAF independently of that theory. We cannot identify the subject matter of CCAF—namely, categories—by pointing to internal categories, which are defined only relative to a given category. And it seems that other proposals also rely on categories or on set theory in the same way.

How do the other theories fare? Let us consider SDG first. SDG is a theory of smooth spaces, and it asserts the existence of a one-dimensional continuum containing infinitesimals, as well as product and function spaces for any pair of spaces. Thus, if it is to avoid the Mismatch Objection, it must be possible to identify its subject matter independently of the theory. And, if it is to avoid the Logical Dependence Objection, it must be possible to do so independently of a prior theory of sets. Traditionally, mathematical spaces are defined to be sets of points equipped with a certain structure—a topological structure, for instance, or a geometric structure. So it might be thought that, while SDG answers the Mismatch Objection, it does so in such a way that it falls foul of the Logical Dependence Objection. However, this

[4]For a similar criticism, see [16].

is not the case. The spaces considered by SDG are not assumed to be mathematical structures composed of sets equipped with functions and relations. Indeed, nothing at all is assumed about their internal composition. All that is assumed is what is contained in the axioms, and this is stated only in terms of mappings between the spaces. In fact, it is a consequence of the axioms that many of the spaces of SDG simply cannot be considered as sets of points: on the most natural understanding of a set of points in SDG, two quite different spaces can have the same set of points ([30]). But, unlike the subject matter of CCAF, the spaces of SDG can be identified independently of the theory. We have a conception of space that is prior to any attempt to axiomatize it. This renders the theory contentful, and saves it from the Mismatch Objection.

Next, let us consider ETCS. This is a theory of sets and, as we will see below, it asserts outright the existence of an empty set, singleton sets, and an infinite set, as well as making hypothetical assertions concerning, for instance, the existence of a power set for any given set. It might seem that this is most obviously vulnerable to the Logical Dependence Objection. However, while it is itself a theory of sets and functions, it does not depend on a *prior* theory of these entities. Rather, it provides such a theory. Just as ZFC cannot be criticized for relying upon a prior theory of sets and functions, nor can ETCS. Similar remarks apply to Makkai's SFAM, which is based upon a categorical theory of sets. It cannot be criticized for depending on a theory of sets, since it is part of the purpose of SFAM to provide such a theory.

For the remainder of the paper, we focus our attention on the theory ETCS, Lawvere's theory of the category of sets. We do not wish to suggest that this is the best candidate for a categorical foundation, nor that it enjoys universal support amongst the categorists. Rather, we wish to use it as a case study. We will describe the theory in the next section, and then we will explore its foundational credentials. We submit that the issues that we raise regarding this particular theory will apply *mutatis mutandis* to the other foundational theories that are not vulnerable to the Mismatch or Logical Dependence Objections: that is, SDG, SFAM, Bénabou's theory of fibered categories, and Bell's theory of local set theories. ETCS is the natural theory to take as a case study, since it is in many ways the most straightforward and least sophisticated theory proposed by the categorists.

2 The theory ETCS

In this section, we describe the theory ETCS whose foundational status we will be investigating.

Before stating its axioms, it is worth observing a fundamental difference between ETCS and the orthodox foundations for mathematics in set theory. The single primitive relation involved in an orthodox set-theoretical foundation is the membership relation, which holds between two sets or between an individual and a set. As a result, existential claims in those foundations tend to be accompanied by a full specification of the members of the set whose existence is asserted. For instance, when we assert the existence of Cartesian products, we say that, for all sets A and B, there is a set $A \times B$ whose members are all and only those ordered pairs whose first member belongs to A and whose second member belongs to B, where ordered pairs are a certain sort of set.

By contrast, in categorical set theory, there is no apparatus by which to assert that the membership relation holds between two sets. Our apparatus allows us to talk only of mappings between sets. This is witnessed by the fact that (working with traditional ZFC as our background theory) ETCS has models with completely different membership structures, and none of these models has a greater claim than any other to be the intended interpretation of the theory. For instance, ETCS has a model in the ordinary cumulative hierarchy of sets, where there is a great deal of overlap between sets, and in which there are many sets that are members of others. But it also has models in which no two sets have the same members, and no set is a member of another. Indeed, ETCS even has models where there is only one set of each cardinality. In these models, the subsets of a set are represented by certain sorts of mappings called *monics*, where the definition of a monic is given purely in terms of arrows: intuitively, in ETCS, we think of a subset of A not as a set whose elements are amongst the elements of A—after all, without a membership relation, we have no way of stating this in ETCS—but as an embedding of a set into A by means of a monic map. None of these models has a greater claim to be the intended model: all interpret the primitive vocabulary of sets and mappings in the intended way. In other words, the axioms of ETCS remain agnostic on all such membership relations: they neither rule them out nor rule them in.

Thus, when we make existence claims in ETCS, we do not assert the existence of a particular set by specifying its members. This approach is not open to us. Rather, we say that there is at least one set that, *together with certain mappings*, fills a particular functional role, where this functional role is specified purely in terms of sets and mappings, and makes no reference to the particular members of the sets. For instance, when we assert the existence of Cartesian products in ETCS, we say that, for all sets A and B, there is at least one set $A \times B$ equipped with mappings $\pi_A : A \times B \to A$ and $\pi_B : A \times B \to B$ that plays the role of a Cartesian product (henceforth, we write such pairs of functions as $\pi_A, \pi_B : A \times B \rightrightarrows A, B$). This particular role, and other relevant ones, will be explained below.

We can sum up this difference in the following motto, which embodies the guiding spirit of category theory: ask not what a mathematical object *is*; ask what it *does*.[5] In orthodox set-theoretic foundations, we make existence claims by asserting the existence of a set and saying exactly what it *is*, i.e., what its members are. In category-theoretic foundations, we make these claims by saying what a set equipped with some mappings needs to *do* to count as a certain sort of object; and we assert that there is at least one object of that sort.

With this difference in mind, we now provide the promised explanation of ETCS.[6] (*Cognoscenti* may consider skipping ahead.) In the terminology of category theory, ETCS says that together the sets and the mappings between them form a non-degenerate, well-pointed topos that contains a natural number object and which satisfies the axiom of choice. We will explain each of these claims in turn. In each case, we first define the relevant category-theoretic terminology and state the axiom using this terminology. All this takes place in the general context of an arbitrary category. Then we give a heuristic explanation of what the axiom says of the universe of sets and the mappings between them. In these heuristic explanations, we will sometimes talk of the elements of a set. Admittedly, as we said above, there is no membership relation in ETCS. But we will see below that it is possible to define a restricted notion of element in ETCS, which suffices to explain and motivate all but the first of our axioms. One final point: We have chosen the present axiomatization for expository purposes. It contains some redundancy, as some of the axioms follow from the others.

A topos is a particular sort of category. So to say that the sets and mappings together form a topos is to say first that they form a category: that is, each mapping is assigned a

[5]In standard category theory, objects are characterized uniquely only up to functional role, whereas mappings are characterized uniquely up to identity. The situation is different in the theory of so-called *2-categories*, or *n-categories* more generally. Regardless, the point remains that objects are only ever characterized in terms of their mapping properties.

[6]We thank an anonymous referee for this journal for detailed comments on the explanation that follows. These helped us improve it greatly.

domain and codomain; the composition-of-mappings operator $\circ$ is associative; and a unique identity mapping $\mathsf{Id}_A : A \to A$ exists for each set A.

To say that the category of sets and mappings is a non-degenerate topos is to make two outright existence claims, three hypothetical existence claims, and a claim about how some of the objects said to exist are related. These claims are expressed by the following five axioms. As is common practice, we will often abbreviate '$f \circ g$' by 'fg' in what follows.

Definition 2.1 *Suppose $\mathcal{C}$ is a category and* 0 *and* 1 *are objects of $\mathcal{C}$. Then*

(i) 0 is an initial object of $\mathcal{C}$ *if, for every A, there is a unique arrow* $0_A : 0 \to A$.

(ii) 1 is a terminal object of $\mathcal{C}$ *if, for every A, there is a unique arrow* $1_A : A \to 1$.

Axiom 1 (Initial and terminal objects) *There is an* initial object *and a* terminal object*.*

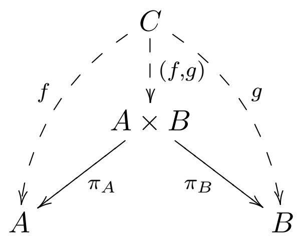
We can easily show that all initial objects are *isomorphic*: that is, for all initial objects 0 and $0'$, there are arrows $f : 0 \to 0'$ and $g : 0' \to 0$ such that $fg = \mathsf{Id}_{0'}$ and $gf = \mathsf{Id}_0$. Similarly, all terminal objects are isomorphic. Henceforth, we pick an arbitrary initial object and denote it 0, and we pick an arbitrary terminal object and denote it 1.

Axiom 2 (Non-degeneracy) 0 *and* 1 *are not isomorphic.*

Using the terminal object introduced by these axioms, we can give the promised explanation of a restricted notion of element that is available in ETCS. The idea is to use maps from the terminal object 1 to express claims that in orthodox set theory are expressed using the membership relation. Heuristically, the idea is that any map from the terminal object into a given set will pick out a single member of that set—a set 1 into which there can be only one function from any given set must be such that any function f from 1 into a given set A will pick out a single member of A. Thus, in ETCS, an *element of* A is a mapping $x : 1 \to A$. Of course, we cannot recover all set-theoretic talk of membership by this device. For instance, there is no way of saying that one set A is an element of another set B, since an element of B, on this definition, isn't a set at all but rather a map from 1 to B. Nor is there any way of saying that an element of one set A is also an element of a different set B. For an element of A is a map from 1 to A, and thus cannot possibly be an element of B, which would be a map from 1 to B. These limitations are often expressed by saying that, in ETCS, there is no *global* membership relation. However, it does allow us to speak of the element of B to which an element of A is taken by a mapping from A to B. For $f : A \to B$ takes an element $x : 1 \to A$ of A to the element $fx : 1 \to B$ of B. This will suffice for all of the talk of elements that we use below.

With this definition of a restricted membership relation, we get from our first axiom that 1 has exactly one element, since there is exactly one mapping $1_1 : 1 \to 1$. Thus, terminal objects are singleton sets. And it follows from the first two axioms together that 0 has no elements. Thus, initial objects are empty sets. In orthodox set theory, the axiom of extensionality guarantees that there is at most one empty set. As we will see below, the version of extensionality that we give in ETCS does not guarantee this.

Definition 2.2 *Suppose $\mathcal{C}$ is a category and A and B are objects of $\mathcal{C}$. Then* $\pi_A, \pi_B : A \times B \rightrightarrows A, B$ is a Cartesian product of A and B *if, for any $f, g : C \rightrightarrows A, B$, there is unique $(f, g) : C \to A \times B$ such that the following diagram commutes:*[7]



Axiom 3 (Cartesian products) *For any two sets A and B, there is a* Cartesian product *$\pi_A, \pi_B : A \times B \rightrightarrows A, B$.*

Thus, in the category of sets and mappings, a Cartesian product of A and B is a set $A \times B$, equipped with mappings $\pi_A, \pi_B : A \times B \rightrightarrows A, B$, that can represent any pair of mappings $f, g : C \rightrightarrows A, B$ uniquely as a single map $(f, g) : C \to A \times B$. That is, for each such pair of mappings f and g, we can recover f by applying π_A to (f, g) and we can recover g by applying π_B to (f, g): that is, $f = \pi_A \circ (f, g)$ and $g = \pi_B \circ (f, g)$, as asserted by the commutative diagram. And (f, g) is the only mapping that has this property.

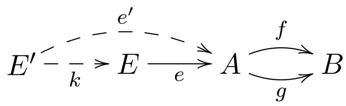
[7]In category theory, the nodes of a commutative diagram represent objects in the category in question, while the edges represent the arrows or mappings. In such a diagram, we usually omit composed mappings and identity mappings. If there are two or more routes through the arrows from one object to another, we say that the diagram commutes if the mappings that result from composing, in order, the arrows that make up these routes are identical. Thus, commutative diagrams are used to make assertions of identity between mappings.

Definition 2.3 *Suppose $\mathcal{C}$ is a category and $f, g : A \rightrightarrows B$ are arrows in $\mathcal{C}$. Then $e : E \to A$* is an equalizer for f and g *if*

(i) $fe = ge$

(ii) *if $e' : E' \to A$ and $fe' = ge'$, then there is a unique $k : E' \to E$ such that $e' = ek$.*

The following commutative diagram illustrates the situation:



Axiom 4 (Equalizers) *For any two mappings $f, g : A \rightrightarrows B$, there is an* equalizer *$e : E \to A$.*

A heuristic explanation: In the category of sets and mappings, an equalizer of f and g is a set E equipped with a mapping $e : E \to A$ such that an element of A is in the range of e if, and only if, f and g agree on that element. Note that this can be rendered entirely in terms of elements conceived as mappings from the terminal object. For instance, to say that an element $a : 1 \to A$ is in the range of e is to say that there is an element $x : 1 \to E$ such that $ex = a$. Thus, although our heuristic explanation talks of elements, it could be written entirely in the language of ETCS. Moreover, so rendered, it is equivalent to the definition just given in the presence of Axiom 6.

As we said above, just as we recover something of the notion of membership in ETCS by talking of maps from the terminal object into a set, we can also recover something of the notion of subset by talking of a certain sort of mapping from one set into another. These mappings are called *monics*:

Definition 2.4 *Suppose $\mathcal{C}$ is a category and $i : B \to A$ is an arrow in $\mathcal{C}$. Then i* is monic *if, for any C and any two distinct arrows $f, g : C \rightrightarrows B$, the arrows if and ig are also distinct.*

In this sense of subset, an equalizer of f and g is intended to represent the subset of A consisting of the elements on which f and g agree. It follows from the definition that if $e : E \to A$ is an equalizer of f and g, then it is a monic mapping into A and thus a subset of A in the sense available in ETCS. Thus, this axiom gives us some of the power of the Axiom

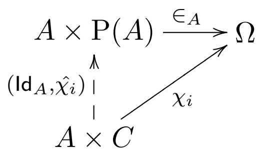
Schema of Subset Separation from orthodox set theory since it guarantees the existence of subsets with certain properties.

In orthodox set theory, Separation guarantees the existence of subsets of a given set with certain properties, while the Axiom of the Power Set guarantees that there is a set that contains all the subsets of a given set. The next axiom of ETCS attempts to capture the same idea. However, in the absence of a global membership relation, we cannot give an exactly analogous statement, since we cannot say that a subset of one set is a member of another. Instead, we say two things. First: Any subset $i : B \to A$ of a given set A can be represented by a unique mapping $\chi_i : A \to \Omega$, where Ω is intended to represent the set of truth-values. This unique function from A to Ω is called the *characteristic function* of the subset, and it takes to the truth-value Truth all and only those elements of A that are in the range of i. Second: Suppose A is a set; then there is a further set $\mathrm{P}(A)$ together with a mapping $\in_A : A \times \mathrm{P}(A) \to \Omega$ such that, for any set C, any subset $i : D \to A \times C$ can be represented uniquely by a mapping $\hat{\chi_i} : C \to \mathrm{P}(A)$, which takes any element c of C to an element of $\mathrm{P}(A)$ that represents the subset of A that contains all and only those elements a of A such that (a, c) is in the range of i. More precisely, (a, c) is an element of the subset D if, and only if, $\in_A$ takes the pair $(a, \hat{\chi_i}(c))$ to the truth value Truth. We call $\in_A$ a *local membership mapping for* A *and* $\mathrm{P}(A)$ since it sends pairs $(a, \hat{\chi_i}(c))$ to Truth if, and only if, a is an element of the subset of A represented by $\hat{\chi_i}(c)$. Combining these two stages: For any set A, the elements of $\mathrm{P}(A)$ represent the subsets of A. As above, this heuristic explanation can be given in the language of ETCS, and, so rendered, it is equivalent in the presence of Axiom 6 to the axiom we now state.

Definition 2.5 *Suppose $\mathcal{C}$ is a category.*

(i) *Suppose $\mathcal{C}$ contains a terminal object* 1*. Then* $\mathsf{true} : 1 \to \Omega$ is a subobject classifier *if, for any monic $i : B \to A$, there is a unique arrow $\chi_i : A \to \Omega$ such that i is an equalizer of* $\chi_i, \mathsf{true} \circ 1_A : A \rightrightarrows \Omega$.

(ii) *Suppose $\mathcal{C}$ contains a subobject classifier* $\mathsf{true} : 1 \to \Omega$*. Then* $\mathrm{P}(A)$ together with $\in_A : A \times \mathrm{P}(A) \to \Omega$ is a power object of A *if, for any monic $i : D \to A \times C$, there is a* *unique* $\hat{\chi}_i : C \to \mathrm{P}(A)$ *such that the diagram below commutes:*



Axiom 5 (Subobject classifier and Power object)

(i) *There is a* subobject classifier $\mathsf{true} : 1 \to \Omega$.

(ii) *For any set* A*, there is a* power object $\mathrm{P}(A)$ *equipped with* $\in_A : A \times \mathrm{P}(A) \to \Omega$.

In ETCS, we call power objects *power sets.* And we call $\in_A : A \times \mathrm{P}(A) \to \Omega$ the *local membership relation* for A.

Axioms 1–5 amount to the assertion that the category of sets and mappings is a non-degenerate topos. But ETCS does not stop there. It goes on to ascribe to that topos various other features.

Axiom 6 (Well-pointedness) *If* $f, g : A \rightrightarrows B$ *and* $fx = gx$ *for all* $x : 1 \to A$*, then* $f = g$.

If, as usual, we represent the elements of a set A by mappings $x : 1 \to A$, then this axiom says that a mapping on a set A is determined solely by its behaviour on the elements of A. Thus, well-pointedness is an extensionality axiom for mappings. However, as noted above, it does not amount to an extensionality axiom for sets, since ETCS has models containing many empty sets.

It is this axiom that allowed us above to state heuristic explanations of the equalizer and power object axioms in terms of elements. Well-pointedness entails that nothing is lost by doing this. In a category with a terminal object in which well-pointedness fails, the heuristic explanations we gave of these two axioms in terms of mappings from the terminal object would not entail the official statement of the axioms.

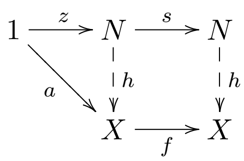
It is a sophisticated, but important result that, together with the axiom of non-degeneracy, well-pointedness entails that Ω must be a two-element set, whose elements represent the truth values Truth and Falsity.

Definition 2.6 *Suppose* $\mathcal{C}$ *is a category and* $j : A \to B$ *is an arrow in* $\mathcal{C}$*. Then* j is epic *if, for any* C *and any two distinct arrows* $g, h : B \rightrightarrows C$*, the arrows* gj *and* hj *are also distinct.*

Axiom 7 (Choice) *If $j : A \to B$ is epic, then there is $g : B \to A$ for which $jg = \mathsf{Id}_B$.*

In the category of sets and mappings, Choice makes a hypothetical existence claim, but it concerns mappings, not sets. This is a faithful statement of the axiom of choice, which in orthodox set theory says that every non-empty set of disjoint non-empty sets has a choice set. In ETCS, we represent a non-empty set of disjoint non-empty sets as a surjective function $j : A \to B$. The disjoint sets are thus represented as the subsets $j^{-1}(b)$ indexed by the elements b of B. The axiom then asserts the existence of a function $g : B \to A$ that picks out a single element of each such disjoint set. As usual, this heuristic explanation can be stated in the language of ETCS.

Definition 2.7 *Suppose $\mathcal{C}$ is a category with a terminal object* 1*. Then N* together with $z : 1 \to N$ and $s : N \to N$ is a natural number object in $\mathcal{C}$ *if, for any object X and arrows $a : 1 \to X$ and $f : X \to X$, there is a unique arrow $h : N \to X$ such that the following diagram commutes:*



Axiom 8 (Natural number object) *The category of sets and mappings contains a natural number object.*

A heuristic explanation: In the category of sets and mappings, a *natural number object* is a set N equipped with two mappings $z : 1 \to N$ and $s : N \to N$ that together guarantee the effectiveness of any recursive definition. That is, for any set X with an initial element picked out by $a : 1 \to X$ and a mapping $f : X \to X$, there is a unique function $h : N \to X$ that takes the zero element of N to a and takes the successor of a 'number' in N to the element of X that results from applying f to whatever element of X was assigned to that number by h: in other words, $hz = a$ and $hs = fh$, as in the commutative diagram above. And again, this explanation may be paraphrased in the language of ETCS, making use of its limited ability to speak to elements.

Notice that this reverses Dedekind's definition of a *simply infinite system.* Dedekind defines a simply infinite system to be a set equipped with an initial element and a successor function such that it is the smallest set containing that element and closed under that successor function. He then proves that such a set guarantees the effectiveness of any recursive definition. By contrast, a natural number object is defined to be something that guarantees the effectiveness of recursion. That it satisfies Dedekind's definition of a simply infinite system is then derived as a theorem.

This completes our presentation of ETCS. The theory is mutually interpretable with the orthodox set theory Z_0C, where Z_0 is Zermelo set theory with subset comprehension axioms only for bounded quantifier formulae, and C is the axiom of choice. As [34] shows, it is possible to introduce natural category-theoretic counterparts of full subset comprehension, as well as full replacement, in order to give extensions of ETCS that are mutually interpretable with ZF and ZFC. The same is true of many of the usual large cardinal axioms (51, [33]).

3 The autonomy of theories

We saw above that ETCS is vulnerable neither to the Mismatch Objection nor to the Logical Dependence Objection. However, a putative foundation for mathematics must boast more than mere logical autonomy with respect to set theory if it is to be truly autonomous. It must be possible not only to *formulate* the foundation without presupposing a theory of sets; it must be possible also to *understand* it and to *justify its claims* without such a presupposition. Unless these further conditions are met, the foundation does not truly support the discipline of mathematics on its own and independent of other assumptions.

Thus, we introduce the distinction between *logical*, *conceptual*, and *justificatory autonomy.* Suppose T_1 and T_2 are theories: not the formal theories of mathematical logic, but rather accounts of a particular part of reality.

- T_1 has *logical autonomy* with respect to T_2 if it is possible to formulate T_1 without appealing to notions that belong to T_2.

- T_1 has *conceptual autonomy* with respect to T_2 if it is possible to understand T_1 without first understanding notions that belong to T_2.

- T_1 has *justificatory autonomy* with respect to T_2 if it is possible to motivate and justify the claims of T_1 without appealing to T_2, or to justifications that belong to T_2.

Such talk of a notion *belonging* to a theory should be tolerably clear already at this point, and it will become clearer when we give examples below.

We saw above that ETCS enjoys logical autonomy with respect to the orthodox foundation for mathematics in set theory. It is possible to state ETCS without appealing to notions that must be introduced by orthodox set theory. Of course, it is not possible to state ETCS without appealing to the notion of set and mapping. But while these notions can be introduced by orthodox set theory, they do not belong specifically to that theory. Rather, ETCS has an equal claim to them.

In the next section, we consider the conceptual autonomy of ETCS with respect to orthodox set theory. After that, we turn to the justificatory autonomy of ETCS, and we consider how the categorist might argue for this and how the orthodox set theorist might respond. As noted above, we do not conclude in favour of one or the other. Our purpose is to explore the terrain.

4 The conceptual autonomy of ETCS

Given that ETCS enjoys logical autonomy with respect to set theory, why might we think that it is not conceptually autonomous? We consider two objections to the conceptual autonomy of ETCS.

The first is due to Dan Isaacson. His complaint is this. ETCS is stated in terms of initial objects, terminal objects, Cartesian products, subobject classifiers, and power objects. These are characterized solely in virtue of their mapping-theoretic properties. Thus, in order to state the corresponding axioms, we need not appeal to the membership relation or other apparatus that belongs to the orthodox set-theoretic foundation. It is for this reason that ETCS is logically autonomous. However, whenever we come to explain these axioms to those unfamiliar with them, we inevitably appeal to the membership relation, the subset relation, the notion of ordered pair, the notion of a function as a set of ordered pairs, and so on. That is, at the point of explanation, the mapping-theoretic presentation is abandoned in favour of a more orthodox presentation, which is required to allow us to understand the axioms. Thus, ETCS does not have conceptual autonomy.

We see two responses to this objection. The first rejects the requirement of conceptual autonomy on the grounds that it is too subjective. Different people with different educational backgrounds will order theories differently with respect to conceptual dependence. What is required for understanding in one individual need not be required in another. So we must abandon our requirement that a foundation be conceptually autonomous.

This response seems partially correct. There will certainly be pairs of theories for which the relation of conceptual dependence is not clear. But this does not rule out the possibility of theories for which this relation is perfectly clear and objective. For instance, whenever T_1 is logically dependent on T_2, then it is objectively the case that T_1 is also conceptually dependent on T_2. Perhaps Isaacson's objection succeeds in showing that ETCS is not conceptually autonomous with respect to set theory in just such an objective sense.

This brings us to the second response to the objection. This response accepts the requirement of conceptual autonomy and argues that ETCS satisfies it. While it is certainly often easier to explain the axioms of ETCS by appealing to their counterparts in orthodox set theory, this is not necessary. Rather, each axiom can be glossed in a way that is quite independent of the membership relation and other apparatus peculiar to orthodox set theory. We attempted such a gloss in section 2; Lawvere and Rosebrugh have attempted a similar project in their introductory text on ETCS ([23]). In these presentations, there is no reference to the sort of membership relation that would allow us to identify members of different sets, or to assert that one set is a member of another. We submit that these introductory glosses are autonomous with respect to any notions that belong peculiarly to orthodox set theory.

To illustrate the claim, consider the notion of Cartesian product. We submit that the following three requirements provide a plausible and completely autonomous account of what we mean by saying that X is a Cartesian product of two given objects A and B. If conceptual analysis is possible at all, this seems to be an instance of it. Firstly, we want there to be projections $\pi_A, \pi_B : X \rightrightarrows A, B$. Secondly, we want X to consist of "independent representations" of A and B, much like a cylinder consists of "independent representations" of a line and a circle. This idea can be expressed as the requirement that any mapping of an object Y to A and B gives rise to a mapping from Y to X; or, more precisely, that any pair of mappings $f, g : Y \rightrightarrows A, B$ factorizes via X and the projections π_A and π_B. Thirdly, we want X to be minimal in the sense that any ordered pair of an element of A and an element of B has a unique representative in X. This idea can be expressed as the requirement that any pair of mappings $i, j : 1 \rightrightarrows A, B$ factorizes uniquely via X and the projections. These three requirements are easily seen to be equivalent to the official definition of Cartesian product, under the assumption of well-pointedness.[8]

Another objection is due to John Mayberry. The fundamental notions of ETCS are the notions of *set* and *mapping*. But according to this objection, the notion of mapping can only be understood by appeal to the orthodox set theorist's reduction of mappings to sets of ordered pairs. Historically, the notion of mapping arises as an idealization of the notion of a rule. And this is also how we introduce it in mathematics education. The objector submits that the only precise account of the notion of mapping that captures the level of idealization that is required in modern mathematics is given by the definition of a function as a set of ordered pairs that represents a many-one or one-one relation, and this definition belongs essentially to orthodox set theory. So in order to understand ETCS, we must appeal at least to this part of orthodox set theory. Thus, ETCS is not conceptually autonomous with respect to the orthodox foundation.

The categorist may respond to this objection as follows. Both sides of the dispute accept that the notion of a mapping precedes the set theorist's reduction of mappings to sets of ordered pairs. The difference is that the set theorist holds that only such a reduction can make the notion sufficiently precise. However, given a basic notion in a particular discipline, there are at least two sorts of account we can give of that notion. We can give a *reductive* (or *explicit*) account, which characterizes the notion in terms of something less problematic; or we can give an *axiomatic* (or *implicit*) account, which characterizes the notion by stating substantial facts in terms of that notion. The set theorist takes the former approach to the notion of mapping; the categorist takes the latter. While a reductive account is usually to be preferred, it doesn't follow that an axiomatic account must depend conceptually upon a reductive one. Thus, the objection is defeated.

5 The justificatory autonomy of ETCS

We have seen that the categorist can plausibly claim logical and conceptual autonomy for her putative foundation for mathematics in ETCS, the categorical theory of sets. We now ask whether she can also claim justificatory autonomy for it.

As we have seen, ETCS is an assertory theory: it makes many existential claims, both

[8] In the much less intuitive case of non-well-pointedness, the third requirement needs to be strengthened.

categorical and hypothetical. So if it is to provide an autonomous foundation for mathematics, it must be able to justify its assertions without appealing to orthodox set theory, or to any aspect of the justification of orthodox set theory that belongs primarily to that theory. So our first task is to consider the justification of orthodox set theory.

5.1 The iterative conception as a justification of ZFC

The standard justification for the axioms of orthodox set theory lies in the iterative conception of set ([12], [9], [35]). According to this conception, the universe of sets may be divided into a well-ordered hierarchy of levels. To each set is assigned a level of this hierarchy in such a way that all elements of that set are assigned to strictly lower levels of the hierarchy. Thus, a set can occupy a stage of the hierarchy only if all of its elements are already present at lower levels. What's more, a set occupies only the lowest level of the hierarchy that it can occupy; it does not recur again at any higher levels.

With this framework in place, the iterative conception of set amounts to the following claim of set-theoretic plenitude: relative to the constraints on the hierarchy just stated, whenever a set *could* occupy a level of the hierarchy, it *does*. To see this in action, consider the power set axiom of Zermelo set theory. Suppose A is a set. Then A occurs at some level λ of the hierarchy. Now suppose X is a subset of A. Since all elements of X are elements of A, they all occur at levels lower than λ. It follows that X must occur at level λ or below. Since X was arbitrary, all subsets of A occur at level λ or below. So at the first level above λ, it is possible for there to exist a set $\mathrm{P}(A)$ that contains all subsets of A. Hence by the plenitude claim there is such a set. In this way, the power set axiom is justified. Similar justifications can be given for the axioms of empty set, pair set, union, subset separation, and foundation. Infinity requires that there be an infinite level of the hierarchy, and replacement requires that the levels of the hierarchy satisfy a cofinality condition. Whether these latter are genuinely extra assumptions in addition to the iterative conception's plenitude claim is a matter of debate, but it need not detain us here. Neither need we consider the vexed question of whether the axiom of choice is justified by appeal to the plenitude claim ([9], [38]).

Of course, the iterative conception does not supply the sort of justification that will convince a sceptical nominalist who demands a justification for the claim that there are any sets at all. But that problem will face all foundations for mathematics that posit entities whose existence the sceptical nominalist doubts. It will face ETCS just as much as the orthodox foundation. So we may bracket this problem. Nonetheless, the iterative conception does provide a justification: on the assumption that there are any sets at all, it justifies many of the particular claims about what sets there are.

5.2 The question sharpened

We claim that the justification provided by the iterative conception of set belongs primarily to orthodox set theory. The iterative conception describes a hierarchy structured in terms of membership relations. So this relation plays an absolutely fundamental role in the iterative conception. This is reflected in the axioms of ZFC, which are stated precisely in terms of the membership relation. In stark contrast, the categorical theory of sets is agnostic about all relations of identity between elements of different sets and about all relations of membership. As we have seen, ETCS describes sets solely in terms of the functional role that they fill. So this approach refrains from all claims about the relation between sets and their elements.[9] We conclude that orthodox ZFC provides a better articulation of the iterative conception than ETCS, and that the justification provided by this conception thus belongs primarily to orthodox set theory rather than to the categorical approach.

It appears that the iterative conception not only makes fundamental use of the membership relation but also draws on a fairly robust metaphysical conception of this relation. For on the iterative conception, sets are composed of their elements in such a way that the former depend metaphysically upon the latter. As Charles Parsons puts it, the iterative conception is "the conception of set as a totality 'constituted' by its elements, so that it stands in some kind of ontological dependence on its elements, but not vice versa" (332, [36]).[10] On this view, the iterative conception describes substantial metaphysical relations between sets. For instance, it supports the modal claim that a set cannot exist unless all of its elements exist as well. If this metaphysical conception of the membership relation can be made out, it will further strengthen our argument that the justification provided by the iterative conception belongs primarily to orthodox ZFC.

Our question about the justificatory autonomy of ETCS thus becomes: Can ETCS provide

[9] Though it seems that mappings can exist only when their domain and codomain exist.

[10] See [39] for similar claims. However, Parsons in the end concludes that this "ontologically richer conception of set" is not needed for the justification of ZFC (137, [37]).

an analogous justification for its particular existential claims that does not depend in an essential way on the iterative conception, which belongs primarily to the orthodox foundation?

In what follows we divide the possible justifications into two classes, depending on how they interpret these existential claims. According to the first sort of justification, each existential assertion of ETCS is to be understood as asserting the existence of a *particular thing*. For instance, the power object axiom asserts, for each set A, the existence of a particular thing, namely the power set of A, where this is understood as a particular object. According to the second sort of justification, each existential assertion of ETCS is to be understood as making a *general* existence claim; that is, a claim that there is at least one object capable of filling the functional role in question. For instance, the power object axiom asserts the existence of some object or other, equipped with a mapping, which is capable of filling the functional role specified by the axiom for the power object. We consider each sort of justification in turn.

5.3 The sets of ETCS are collections of *lauter Einsen*

On the first sort of justification, the axioms of ETCS assert the existence of particular objects. But in order to remain autonomous with respect to orthodox set theory, these particular objects must be different from the objects described by the iterative conception of set.

Just such an account is given by Lawvere, who first formulated ETCS ([22]). Lawvere describes ETCS as the theory of *abstract sets* and *arbitrary mappings* between them. For Lawvere, abstract sets are quite different from the sets introduced by the iterative conception. Most importantly, while the sets that populate the iterative hierarchy generally have members with a great deal of internal structure given in terms of the membership relation, and a large number of intrinsic properties, Lawvere's abstract sets are collections of what he calls "*lauter Einsen*" or "pure units", following Cantor. That is, the elements of the sets of ETCS have no internal structure and no intrinsic properties, and their distinctness one from another is a brute fact that is not reducible to a fact about distinguishing properties.

One way to view this proposal is to compare it to recent accounts of the abstract entities postulated by *ante rem* (or *sui generis*) structuralism. Just as for Shapiro each natural number has no properties other than those it has in virtue of its position in the natural number structure, the members of Lawvere's abstract sets have no properties other than those they have in virtue of supporting the mappings posited by ETCS. To support these mappings, the facts of their identity and distinctness are crucial, while any further properties are extraneous; thus, they do not possess them.

Viewed in this way, it is no surprise to find that similar conceptions of abstract sets have been given before. The mathematical numbers posited by Plato and Speusippus and rejected by Aristotle are abstract sets in this sense (Aristotle *Metaphysics* XIII), as are the abstract cardinal structures considered by [40], and the edgeless graphs discussed by [24].

Are such 'purely structural' objects metaphysically problematic? ([13], [27], [25]). While this question demands discussion, we restrict ourselves here to the epistemological question of whether such a conception can endow ETCS with justificatory autonomy. Given the understanding of the axioms of ETCS as concerned with a universe of abstract sets of featureless elements, together with the mappings between them, how might the categorist justify those axioms? For instance, how might she justify the claim that for every abstract set A, there is another abstract set $\mathrm{P}(A)$, equipped with a membership mapping $\in_A \colon A \times \mathrm{P}(A) \to \Omega$, that fills the functional role required of a power object for that set? We consider two attempted justifications. Our purpose is not to consider only justifications that have actually been given; rather, we wish to explore the possible moves that could be made.

The first justification is Hilbertian. According to this justification, any consistent theory of a system of abstract objects is true. That is, for any such theory there are abstract objects that answer to the description given by that theory. Thus, to the extent that we are justified in believing ETCS to be consistent, we are also justified in believing that there are abstract sets and arbitrary mappings that it describes.

This justification faces the usual problems that such Hilbertian accounts face. Why, for instance, should we think that consistency entails existence? But even if this question can be answered, a further worry lingers. It is often said that we are justified in believing in the consistency of our mathematical theories because they have been in use for so long, yet have yielded no contradictions. But this is false. It would only be true if contradictions had actively been sought in the places where they are most likely, namely in those parts of the theories that lie closest to paradox. But they haven't. Rather, our confidence in the consistency of arithmetic, real analysis, functional analysis, and even higher set theory is justified on the basis of our clear conception of what the universe of those disciplines is like. The iterative conception of set equips us with an understanding of the structure of the set-theoretic universe that justifies our belief in the consistency of the theory that describes it. And the consistency of the other disciplines follows from this, or from analogous conceptions of their own universes. Thus, by abandoning the iterative conception of set, Lawvere does not just abandon a particular metaphysical picture; he also dismisses a conception of the universe of sets that is crucially involved in our best justification of the consistency of our foundation.

The second justification of the axioms of ETCS as assertions about abstract sets is naturalistic, in the sense that it defers to the opinions of working scientists. Since the assertions of working mathematicians entail the existence of Cartesian products, power sets, infinite sets, and so on, the rest of us too are justified in believing in the existence of these mathematical objects. Alternatively, if one is reluctant to say that we are justified in believing all the assertions of *mathematicians*, the foregoing argument is easily transformed into an indispensability argument, which requires only that we are justified in believing all the assertions of our *current best theory of the physical universe*. After all, as usually formulated, our current best physical theories entail the existence of Cartesian products, power sets, infinite sets and so on. Since we are justified in believing these theories, we are also justified in believing in the existence of these mathematical objects.

Again, these justifications face problems. We now describe the most pressing one. (A more general concern about the idea of a naturalistic justification for ETCS will be developed below.) As has often been observed, the ontology that might be justified by an indispensability argument is underdetermined, since the same successful physical theory might be formulated using different mathematical theories that describe different mathematical objects. In the case of the naturalistic justification of ETCS, even the mathematical practice underdetermines the ontology that might be justified by appealing to it. After all, while it is certainly true that mathematicians assert the existence of Cartesian products, power sets, infinite sets, and so on, they do not say enough about the nature of these objects to determine whether they are the power sets from the iterative conception of sets, or the power sets from Lawvere's conception of abstract sets of featureless elements, or some other sort of power sets. So the prospects for a naturalistic justification for ETCS, interpreted as a theory of Lawvere's abstract sets, seem bleak.

In light of the above comparison with *ante rem* structuralism, it might seem that the following riposte is available to the categorist who wishes to defend Lawvere's conception on naturalist grounds. Recently, Shapiro has argued from mathematical practice to the existence of his *ante rem* structures on the basis of two theses, which he dubs *faithfulness* and *minimalism* (110, [41]). According to the former, we should assert the existence of a mathematical object when and only when the mathematician does; according to the latter, we should not ascribe to these objects any property that the mathematician does not ascribe to them. Shapiro submits that mathematicians ascribe to their objects no properties other than those that the objects have in virtue of belonging to a system with a particular structure. Assuming this claim, it follows from *minimalism* that we should ascribe to mathematical objects only their purely structural properties.

However, this claim is problematic. By asserting positively that the elements of the sets of mathematics are featureless, we are ascribing to them a property that the mathematicians never postulated, namely their featurelessness. Thus, minimalism does not entail that the sets assumed by mathematicians are sets of featureless elements. This undermines both Shapiro's original argument and any attempt to deploy it in defence of Lawvere.

Another objection to the naturalist's justification of ETCS is this: It is simply false that working mathematicians are agnostic about the internal constitution of the sets about which they speak. After all, many textbooks that introduce elementary areas of mathematics, such as algebra, analysis, and number theory, include an introductory section surveying the elements of set theory, and this set theory is explicitly orthodox set theory—in particular, it includes assertions about membership relations that cannot be made in ETCS. Thus, the naturalist proponent of ETCS will have to say either that such assertions are not to be included in the evidence gleaned from mathematical practice, or that such brief introductory assertions are somehow outweighed by the vast majority of mathematical literature that does not reveal commitment to orthodox set theory.

In sum, while Lawvere presents a novel conception of the foundations of mathematics in a theory of abstract sets of pure units, and the mappings between them, it is not clear that it can be used to give a justification of ETCS that is autonomous with respect to the orthodox foundation in set theory. This concludes our discussion of those attempts to justify ETCS that interpret its existential claims as concerning particular entities.

5.4 The sets of ETCS can be just what they have to

We turn finally to the putative justifications of ETCS that interpret its existential claims as general existence claims. For instance, on the interpretation that underlies these justifications, the power set axiom does not assert, for each abstract set of pure units, the existence of a further abstract set of pure units that fills the functional role required of a power set. Rather, it remains agnostic about the nature of the sets with which ETCS is concerned, and merely asserts the existence of *some object* that, together with *some map*, fills the role. Echoing McLarty's claim about the natural numbers conceived category-theoretically, the sets can be "just what they have to" ([32]).

Interpreted thus, how might one justify ETCS? Again, the Hilbertian option and the naturalistic option are open to us. We have nothing to add to our discussion of the putative Hilbertian justification.

However, in the case of the naturalist justification, the situation has changed markedly. Above we objected to the naturalistic justification for ETCS, interpreted as a theory of Lawvere's abstract sets, on the grounds that this interpretation goes beyond what is warranted by mathematicians' own assertions about the sets with which they are concerned. Clearly such an objection cannot be raised against a naturalistic justification of ETCS when this theory is interpreted as making only general existential claims. On the contrary, it seems that ETCS, interpreted in this way, is highly appropriate to the naturalistic argument. After all, if mathematicians remain agnostic about the internal constitution of the objects of their study, then naturalism can at best justify a foundational theory that is similarly agnostic. In other words, if the internal constitution of mathematical objects is not described by working mathematicians, then naturalism will lead to a foundational theory that characterizes its objects only up to isomorphism. And when interpreted in the way under consideration, ETCS is exactly such a theory.

So we submit that naturalism provides the greatest hope for the categorist. If one favours a foundation that respects the non-foundational assertions of working mathematicians, who tend to be agnostic about the internal constitution of their objects, one ought to prefer a foundation that specifies its objects purely in terms of what they do, rather than in terms of what they are: that is, a foundation that specifies its objects only by their functional role, which typically determines an isomorphism class, and not by their intrinsic nature. As we explained above, category theory is ideally suited to such a purpose.

However, any naturalistic justification for ETCS will require a very strong form of naturalism. *Moderate naturalists* about a particular scientific discipline hold that the opinions of scientists working in that discipline can suffice to establish that *there exists* a justification for some philosophically significant claim. But moderate naturalists also recognize the need to identify and articulate the justification that is said to exist within the relevant discipline. For instance, a moderate naturalist about mathematics might take the opinions of mathematicians to establish that there is a justification for the existential claims of traditional, membership-based set theory. However, she will not rest content at this point but will proceed to search for that justification within mathematics itself, perhaps aided by professional mathematicians. And, in the iterative conception of set, she may take herself to have found it. By contrast, *extreme naturalists* claim that the very existence of the opinions of working scientists by itself *provides* the required justification for the claim. No further justification is needed beyond the fact that competent scientists with the relevant expertise assent to the claim in question.

The naturalistic justification for ETCS that we outlined above relies on extreme naturalism. All the justification does is appeal to the opinions that prevail among working mathematicians. Any attempt to articulate some substantive justification for ETCS within mathematics itself would go beyond the naturalistic justification that we outlined. But this also means that any such attempt is likely to compromise the agnosticism that appeared to make ETCS so attractive. For instance, if the substantive justification is the one provided by the iterative conception, then the justification will be better captured by orthodox membership-based set theory.

6 Conclusion

In the first half of the paper, we argued that a number of proposed category-theoretic foundations for mathematics in fact avoid the objections raised by Feferman and expanded by Hellman. In the second half, we focussed our attention of ETCS. Our strategy has been to use this theory as a case study: the problems that arise for ETCS will also arise for the other candidate foundations in category theory. We argued that ETCS enjoys conceptual autonomy, as well as logical autonomy. But the question of justificatory autonomy is harder.

The justificatory autonomy of ETCS depends on what sorts of justification one is willing to accept. Suppose one agrees with the extreme naturalist that it suffices for the justification of ETCS that mathematicians make assertions whose truth requires the existence of things that play the functional roles of power objects, Cartesian products, infinite sets, and so on; that is, that mathematicians specify their foundational objects at most up to isomorphism. Then this will be a justification for ETCS that does not depend on orthodox, membership-based set theory, nor on any justifications that belong primarily to that theory, such as the iterative conception. This will establish that ETCS has justificatory autonomy with respect to orthodox set theory. On the other hand, if one requires that justifications be more substantive than those provided by extreme naturalism, then it seems doubtful that ETCS will have justificatory autonomy.

One final point: Suppose we agree with the extreme naturalist and conclude that ETCS has justificatory autonomy with respect to orthodox set theory. It does not follow that we must also hold that the justification given for orthodox set theory via the iterative conception of set, and the autonomous justification given for ETCS via extreme naturalism are equally good justifications. It is quite consistent to hold that both theories are justified, that each has a justification that is independent of the other, but nonetheless that orthodox set theory is better justified than its category-theoretic counterpart.

References

[1] Steve Awodey. Structure in Mathematics and Logic: A categorical perspective. *Philosophia Mathematica*, 4(3):209–237, 1996.

[2] Steve Awodey. An Answer to Hellman's Question: 'Does Category Theory Provide a Framework for Mathematical Structuralism?'. *Philosophia Mathematica*, 12(3):54–64, 2004.

[3] Steve Awodey. A Brief Introduction to Algebraic Set Theory. *Bulletin of Symbolic Logic*, 14(3):281–298, 2008.

[4] John L. Bell. Category theory and the foundations of mathematics. *British Journal for the Philosophy of Science*, 32:349–358, 1981.

[5] John L. Bell. Categories, toposes and sets. *Synthese*, 51(3):293–337, 1982.

[6] John L. Bell. *Toposes and Local Set Theories*, volume 14 of *Oxford Logic Guides*. Clarendon Press, Oxford, 1988.

[7] J. Bénabou. Fibered Categories and the Foundations of Naive Category Theory. *Journal of Symbolic Logic*, 50(1):10–37, 1985.

[8] Paul Benacerraf and Hilary Putnam, editors. *Philosophy of Mathematics: Selected Readings*. Cambridge University Press, Cambridge, 1983.

[9] George Boolos. The Iterative Conception of Set. *Journal of Philosophy*, 68(8):215–231, 1971.

[10] Solomon Feferman. Categorical Foundations and Foundations of Category Theory. In R. E. Butts and J. Hintikka, editors, *Logic, Foundations of Mathematics and Computability Theory*, pages 149–169. Reidel, Dordrecht, 1977.

[11] Peter Freyd. *Abelian categories*. Harper and Row, New York, 1964.

[12] Kurt Gödel. What is Cantor's Continuum Problem? (revised version). In Benacerraf and Putnam [8], pages 470–485.

[13] Geoffrey Hellman. Three Varieties of Mathematical Structuralism. *Philosophia Mathematica*, 9(3):184–211, 2001.

[14] Geoffrey Hellman. Does Category Theory Provide a Framework for Mathematical Structuralism? *Philosophia Mathematica*, 11(3):129–157, 2003.

[15] Geoffrey Hellman. What Is Categorical Structuralism? In J. van Benthem, G. Heinzmann, M. Rebuschi, and H. Visser, editors, *The Age of Alternative Logics: Assessing philosophy of logic and mathematics today*. Kluwer, Dordrecht, 2006.

[16] Geoffrey Hellman and John L. Bell. Pluralism and the Foundations of Mathematics. In Stephen J. Kellert, Helen E. Longino, and C. Kenneth Waters, editors, *Scientific Pluralism*. University of Minnesota Press, 2006.

[17] Martin Hyland. The Effective Topos. In A. S. Troelstra and Dirk van Dalen, editors, *The L.E.J. Brouwer Centenary Symposium (Noordwijkerhout, 1981)*, Studies in Logic and the Foundations of Mathematics, pages 165–216, Amsterdam, 1982. North-Holland.

[18] Anders Kock. *Synthetic Differential Geometry*. Cambridge University Press, Cambridge, 1981.

[19] F. W. Lawvere. An elementary theory of the category of sets. *Proceedings of the National Academy of Sciences of the U. S. A.*, 52:1506–1511, 1964.

[20] F. W. Lawvere. The category of categories as a foundation for mathematics. In Samuel Eilenberg, editor, *Proceedings of the Conference on Categorical Algebra, La Jolla, 1965*, pages 1–21. Springer-Verlag, Berlin, 1966.

[21] F. W. Lawvere. Categorical dynamics. In *Topos Theoretic Methods in Geometry*, volume 30 of *Various Publications*, pages 1–28. Aarhus University Press, Aarhus, 1979.

[22] F. W. Lawvere. Cohesive Toposes and Cantor's '*lauter Einsen*'. *Philosophia Mathematica*, 2(3):5–15, 1994.

[23] F. W. Lawvere and Robert Rosebrugh. *Sets for Mathematics*. Cambridge University Press, Cambridge, 2003.

[24] Hannes Leitgeb and James Ladyman. Criteria of Identity and Structuralist Ontology. *Philosophia Mathematica*, 16(3):388–396, 2008.

[25] Øystein Linnebo. Structuralism and the Notion of Dependence. *Philosophical Quarterly*, 58:59–79, 2008.

[26] Saunders Mac Lane. *Mathematics: Form and Function*. Springer-Verlag, New York, 1986.

[27] Fraser MacBride. Structuralism Reconsidered. In Stewart Shapiro, editor, *Oxford Handbook of Philosophy of Mathematics and Logic*, pages 563–589. Clarendon Press, Oxford, 2005.

[28] M. Makkai. Towards a Categorical Foundation of Mathematics. In J. A. Makowsky and E. V. Ravve, editors, *Logic Colloquium '95: Proceedings of the Annual European Summer Meeting of the Association of Symbolic Logic, held in Haifa, Israel, August 9-18, 1995*, pages 153–90, Berlin, 1998. Springer-Verlag.

[29] John P. Mayberry. On the Consistency Problem for Set Theory: An Essay on the Cantorian Foundations of Classical Mathematics (I). *British Journal for the Philosophy of Science*, 28(1):1–34, 1977.

[30] Colin McLarty. Defining Sets as Sets of Points. *Journal of Philosophical Logic*, 17:75–90, 1988.

[31] Colin McLarty. Axiomatizing a category of categories. *Journal of Symbolic Logic*, 56:1243–1260, 1991.

[32] Colin McLarty. Numbers Can Be Just What They Have To. *Noûs*, 27(4):487–498, 1993.

[33] Colin McLarty. Exploring Categorical Structuralism. *Philosophia Mathematica*, 12(3):37–53, 2004.

[34] G. Osius. Categorical set theory: A characterization of the category of sets. *Journal of Pure and Applied Algebra*, 4:79–119, 1974.

[35] Charles Parsons. What is the Iterative Conception of Set? In Benacerraf and Putnam [8], pages 503–529.

[36] Charles Parsons. The Structuralist View of Mathematical Objects. *Synthese*, 84(3):303–347, 1990.

[37] Charles Parsons. *Mathematical Thought and Its Objects*. Cambridge University Press, Cambridge, 2008.

[38] Alexander Paseau. Boolos on the Justification of Set Theory. *Philosophia Mathematica*, 15(3):30–53, 2007.

[39] Michael Potter. *Set Theory and its Philosophy: a critical introduction*. Oxford University Press, Oxford, 2004.

[40] Stewart Shapiro. *Philosophy of Mathematics: Structure and Ontology.* Oxford University Press, Oxford, 1997.

[41] Stewart Shapiro. Structure and Identity. In Fraser MacBride, editor, *Identity and Modality*, pages 34–69. Oxford University Press, Oxford, 2006.