We hold these truths to be self-evident: But what do we mean by that?

Stewart Shapiro
The Ohio State University
Arché Research Centre, University of St. Andrews
shapiro.4@osu.edu

At the beginning of *Die Grundlagen der Arithmetik* (§2) [1844], Gottlob Frege observes that "it is in the nature of mathematics to prefer proof, where proof is possible", noting that "Euclid gives proofs of many things which anyone would concede him without question". Frege sets himself the task of providing proofs of such basic arithmetic propositions as "every natural number has a successor", the induction principle, and Kant's bugbear "7+5=12".

Frege's observation was true in Euclid's day, and it remains true now. We all admire the achievements of Euclid, Archimedes, Cauchy, Weierstrass, Dedekind, Frege, and a host of others on providing rigorous proofs of "many things that formerly passed as self-evident", as Frege put it (§1). Many of these are propositions that no one in their right mind would doubt—unless it be on skeptical or nominalist grounds (in which case mathematical proof would not settle the issues).

Nevertheless, thinkers differ widely on *why* it is that we prefer proof, and this dispute goes to the very heart of mathematics. My topic here is closely related to this. It is a commonplace that one cannot prove everything (except perhaps in the trivial sense of a one-line "proof" of an axiom). Frege's observation is that mathematics prefers proof, "*where proof is possible*". What about cases where proof is not possible? What is the epistemic status of the axioms, or basic truths, from which other propositions are derived? If we claim to know the theorems, on the basis of the proofs, then surely we must claim to know the axioms? How?

1 Axioms as definitions: they tell us what we are talking about. A relatively recent perspective is that the axioms of a given branch of mathematics serve as an implicit definition of the primitive terms of the branch. David Hilbert's *Grundlagen der Geometrie* [1899] represents the culmination of a trend toward this model-theoretic perspective on branches of mathematics. The early pages of the book contain phrases like "the axioms of this group define the idea expressed by the word 'between' . . ." and "the axioms of this group define the notion of congruence or motion". The idea is summed up as follows:

> We think of . . . points, straight lines, and planes as having certain mutual relations, which we indicate by means of such words as "are situated", "between", "parallel", "congruent", "continuous", etc. The complete and exact description of these relations follows as a consequence of the *axioms of geometry*.

The central aspect of the axiomatization is that *anything at all* can play the role of the undefined primitives of points, lines, planes, etc., so long as the axioms are satisfied. Hilbert was not out to capture the essence of a specific chunk of reality, be it space, the forms of intuition, or anything else. Otto Blumenthal reports that in a discussion in a Berlin train station in 1891, Hilbert said that in an axiomatization of geometry, "one must always be able to say, instead of 'points, straight lines, and planes', 'tables, chairs, and beer mugs'."[1]

Paul Bernays [1967, 497] summed up the aims of the new geometry:

> A main feature of Hilbert's axiomatization of geometry is that the axiomatic method is presented and practiced in the spirit of the abstract conception of mathematics that arose at the end of the nineteenth century and which has generally been adopted in modern mathematics. It consists in abstracting from the intuitive meaning of the terms . . . and in understanding the assertions (theorems) of the axiomatized theory in a hypothetical sense, that is, as holding true for any interpretation . . . for which the axioms are satisfied. Thus, an axiom system is regarded not as a system of statements about a subject matter but as a system of conditions for what might be called a relational structure . . . [On] this conception of axiomatics, . . . logical reasoning on the basis of the axioms is used not merely as a means of assisting intuition in the study of spatial figures; rather logical dependencies are considered for their own sake, and it is insisted that in reasoning we should rely only on those properties of a figure that either are explicitly assumed or follow logically from the assumptions and axioms.

[1]"Lebensgeschichte" in Hilbert [1935, 388-429]; the story is related on p. 403.

Frege had a more traditional orientation toward axioms. In a letter dated December 27, 1899, Frege tried to correct Hilbert on the nature of definitions and axioms.[2] According to Frege, axioms should express truths; definitions should give the meanings and fix the denotations of terms. These fundamentally different enterprises should not be confused. He remarked that, despite what Hilbert [1899] claims, the book does not provide a definition of, say, "between" since the axiomatization "does not give a characteristic mark by which one could recognize whether the relation Between obtains":

> . . . the meanings of the words "point", "line", "between" are not given, but are assumed to be known in advance . . . [I]t is also left unclear what you call a point. One first thinks of points in the sense of Euclidean geometry, . . . But afterwards you think of a pair of numbers as a point . . . Here the axioms are made to carry a burden that belongs to definitions . . . [B]eside the old meaning of the word "axiom", . . . there emerges another meaning but one which I cannot grasp.

According to Frege, definitions are in sharp contrast with axioms and theorems. The latter

> . . . must not contain a word or sign whose sense and meaning, or whose contribution to the expression of a thought, was not already completely laid down, so that there is no doubt about the sense of the proposition and the thought it expresses. The only question can be whether this thought is true and what its truth rests on. Thus axioms and theorems can never try to lay down the meaning of a sign or word that occurs in them, but it must already be laid down.

According to Frege, if the terms in the proposed "axioms" do not have a prior meaning, then the sentences cannot be true (or false), and thus they cannot be axioms. If the terms do have meaning beforehand, then the "axioms" cannot be definitions.

The crucial Fregean principle here is that every well-formed sentence in a mathematical theory makes a fixed assertion about a fixed collection of objects and concepts. Each such sentence has a truth value determined by the nature of the referenced objects and concepts. Geometry is about (Euclidean) *space*; arithmetic is about *numbers*. Alberto Coffa [1991] calls this "propositionalism". For geometry,

---

[2]The correspondence between Frege and Hilbert is published in Frege [1976] and translated in Frege [1980]. See Blanchette [1996] for an insightful discussion of Frege's notion of logical consequence.

and certain other mathematical theories, Hilbert rejected propositionalism. For him, the theorems of Euclidean geometry do not express propositions, in this sense.

Following a suggestion of Geoffrey Hellman's, let us say that for Frege, the axioms and theorems of arithmetic and geometry are ***assertory***. For Hilbert, axioms and theorems are ***algebraic***. Assertory sentences are meant to express propositions with fixed truth values, while algebraic sentences are schematic, applying to any system of objects that meets the conditions specified by the axioms.

So the algebraist has an answer to our question concerning the epistemic status of axioms: they are definitions, and so not asserted as such. So there is no serious question as to how we know them. Strictly speaking, we do not know them, and we do not know the theorems. What we know, it seems, is that the theorems follow from the axioms.

It is not clear that one can take a Hilbert-style model-theoretic perspective for all of mathematics. Consider, for example the following statement:

(*) In any Euclidean system *S*, for any two distinct objects ***a***, ***b*** in *S* that are "points-in-*S*", there is a unique "line-in-*S*" on which ***a*** and ***b*** lie.

For the algebraist, (*) is a statement that a given theorem follows from a given set of axioms. Prima facie, (*) is assertory, telling us something about the relation of logical consequence. It also seems to be a statement within mathematics. Can we understand (*) as algebraic? How? What are the axioms for the theory in which (*) is a theorem? And there would be a (*)-like statement for such a theory. What are we to make of it?

Along similar lines, Hilbert's *Grundlagen* provided consistency and independence proofs of various sets of axioms by finding interpretations that satisfy them. Typically, he would interpret the axioms of a theory in terms of constructions on real numbers. This approach is now as common as anything in mathematics. This free reinterpretation of axioms is a main strength of contemporary mathematical logic, and a mainstay of mathematics generally. What is the algebraist to make of the statements of consistency themselves? What, for example, is the epistemic status of the statement that Euclidean geometry is consistent? Are such statements algebraic or assertory, or both at once?

It seems clear that for Hilbert and just about anyone else, consistency is itself a mathematical matter. His methodology indicates that in order for us to be assured that certain mathematical objects exist, we have to *establish* the consistency, or perhaps the satisfiability, of an axiomatization. In the *Grundlagen*, Hilbert discharged this burden, at least in part, by providing relative consistency *proofs*. For example, he showed how to interpret the axioms of a non-Euclidean geometry in the real numbers, thus establishing that the non-Euclidean theory is consistent if real analysis is.

We now enter the realm of meta-mathematics. Given the way this matter is handled in Hilbert's *Grundlagen*, it is clear that meta-mathematics is itself mathematics. What are we to make of it? What is the status of *statements* and proofs of consistency? Are they assertory or algebraic?

This matter is not treated explicitly in Hilbert's *Grundlagen*, and it is hard to be definitive on what his view was, or should have been, but I suggest that the meta-theory—the mathematical theory in which the consistency of an axiomatization is established—is not to be understood algebraically, not as another theory of whatever satisfies *its* axioms. Instead, the statement that a given theory, such as Euclidean geometry, is consistent is itself assertory. Indeed, Hilbert asserts it, or at least seems to. The notion of consistency is a contentful property of theories, and is not to be understood as defined implicitly by the axioms of the meta-theory. For one thing, the meta-theory is not axiomatized in the *Grundlagen*, and so there is no implicit definition of the meta-theoretic notions. This, of course, is not decisive. It would be a routine exercise to axiomatize the meta-theory of the *Grundlagen*. More importantly, an algebraist who thinks of the axiomatized meta-theory as algebraic would have to worry about *its* consistency. How would we establish that? The ensuing regress is vicious to the epistemological goals of the *Grundlagen*.[3]

In the later Hilbert program (e.g., [1925]) relative consistency gives way to absolute consistency. And this, too, is a mathematical matter. Consistency is to be *proved*. The envisioned meta-theory for this is finitary proof theory, focused directly on formal languages themselves. And Hilbert is explicit that finitary proof theory is contentful, and thus assertory. He was thus a propositionalist about this proof theory. Proof theory should be axiomatized, and, once it is, we can ask our question about the epistemic status of its axioms, or fundamental propositions.

Of course, finitary proof theory proved to be all but useless for establishing consistency, thanks to Kurt Gödel's incompleteness theorem. But the powerful model-theoretic approach to mathematics continues to thrive today, as well it should. But, in practice, how do we satisfy ourselves that a given characterization—whether it is a traditional axiomatization or a type of category—is coherent, and thus characterizes a structure or a possible system?

On the contemporary scene, questions of coherence, satisfiability, and the like, are referred to set theory, which has become the default meta-mathematics. Prima facie, to play this role of helping to decide questions about how the theories relate to each other, and whether a given theory is consistent or satisfiable, set theory must be assertory. That is, we need to *assert* that a given theory is consistent, satisfiable, or whatever, and it is the language of set theory in which these assertions are made, and negotiated. Thus construed, set theory is not just a theory of whatever satisfies its axioms. And so our question about the epistemic status of its axioms is still pressing. So the Hilbertian algebraist does not completely avoid the question of this paper (see Shapiro [2005]).

---

[3] I am indebted to an anonymous referee for *Philosophia Mathematica* for this point.

2 It is obvious, my dear Watson. The traditional view is that the axioms of a properly formulated (assertory) mathematical theory should be *self-evident*, and the proofs should proceed by self-evident steps. This *Euclidean* approach puts the theorems on the most solid epistemic foundation possible: they cannot be more secure, or more certain. It is not much of a distortion to describe rationalism as the attempt to extend this Euclidean methodology to all knowledge, or at least all of science.

So what of the starting point? What is this notion of self-evidence? Famously, Gödel [1963, 484] remarked that the axioms of set theory "force themselves upon us as being true". I do not wish to engage exegetical issues here, but this at least sounds like a statement of the psychology, or perhaps the phenomenology, of mathematics, or at least of set theory. The axioms are *obvious*. Write them down, think about what they say, and you will come to believe that they are true.

Note, however, that at least some of the axioms of set theory do not force themselves on us as true on a first or even a second or third reading. The axioms of infinity and replacement, for example, are hardly obvious at first. If these axioms force themselves on us, it is only after we become familiar with the system, and perhaps its intuitive underpinning, the iterative conception of set. We work with the language, derive some basic theorems, develop intuitions, check the theorems against these intuitions, etc. In other words, we immerse ourselves in the practice. Then, perhaps, after a while it all becomes obvious. Then, perhaps, the axioms force themselves on us.

Even at that point, there is room for doubt due to the extensive ontological consequences of the axioms. The axioms entail the existence of a cardinal κ that is a fixed point in the series of alephs: $\aleph_\kappa=\kappa$. As Boolos [1998] asked, do we really believe in the existence of such a set?[4] To taken another example,

---

[4]The first fixed point in the aleph series might not be all that large. It is consistent with ZFC that κ is smaller than the continuum. Boolos could have picked another example, perhaps a fixed point in the beth series: a cardinal λ such that $\lambda=\beth_\lambda$. Even this is not large by set-theorist's standards, but it is large enough to boggle the mind.

do we really have to believe in the existence of a decomposition of a unit sphere that, recombined via rigid motions, yields two unit spheres? Does this at least shed some doubt on the axioms, or at least on their status as obvious?

In any case, we are asking here about the *epistemic* status of the axioms of assertory mathematical theories, not the psychological states of some or even typical practitioners of the theory. Presumably, the theorems are known on the basis of the axioms (and the logic). How are the axioms *known*? Whether obviousness, or obviousness-after-immersion, has epistemic weight depends on what the correct epistemology for mathematics is. Moreover, if it immersion on the practice that makes the axioms obvious, then the axioms are not properly foundational. Some of the theorems play a role in how the axioms become obvious—how they force themselves on us. It is the practice as a whole that generates the psychological state or the phenomenological feel of certainty.

From this perspective, what becomes of the slogan that "it is in the nature of mathematics to prefer proof, where proof is possible"? The reason for this, it seems, is to trace logical connections among various propositions within a given mathematical theory. This is part of what it takes to immerse oneself in a practice, which is what it takes for the axioms to force themselves on us as true. The more proofs we have, the stronger the feeling of certainty. But, at this level, it is still merely a feeling.

3 Frege: objective self-evidence, epistemic foundation. The typical Euclidean demands a better foundation than this. The idea is that the axioms are *epistemically* secure, and not (merely) obvious. We know the axioms, individually, and this knowledge does not rely on anything else. The rest of the theory is built on the foundation of the axioms.

The quintessential rationalist, Renee Descartes, said that foundational propositions become known once when we have clear and distinct ideas of the concepts that figure in them. Coffa [1991, 10] quipped that despite (or because of) a voluminous literature on this topic, the notion a “clear and distinct idea” is heterological. It is neither clear nor distinct what it is to be or have a clear and distinct idea of something. Presumably, one gets a clear and distinct idea, and thus foundational knowledge of an axiom, by fully coming to understand the concepts involved in the axiom. One can understand and perhaps even know the constituents of a proposition in a less then clear and distinct manner, but in that case, one does not have the appropriate foundational knowledge of it. Prima facie, no holistic considerations are involved in coming to fully understand a potential axiom. It is not a matter of immersion in the practice; rather, we properly reflect on the concepts invoked in the axiom. One this is accomplished, we know the axiom in the proper manner.

It is not clear how tenable this package is for contemporary mathematical theories. I do not see how one can maintain that the axioms of choice, infinity, and replacement can be known on the basis of clear and distinct ideas of the one concept involved, namely set-membership, or even in terms of the intuitive iterative conception. Even more elementary mathematical theories are suspect. Do we have such foundational knowledge of the principles that zero exists and that every natural number has a successor? Do we have foundational knowledge that the universe is infinite? How did we get such knowledge?

A rationalist might argue that the aforementioned principles of set theory and arithmetic are not foundational. Perhaps they can be proved, on the basis of other, more basic primitive propositions, and perhaps we can obtain clear and distinct ideas of those more basic propositions. From this perspective, one of the reasons why “we prefer proof, where proof is possible” is to uncover the foundations of our knowledge. The burden is to provide proper foundational proofs of the basic principles of (assertory) mathematical theories, and then show that the basic principles of such proofs are known, once we have clear and distinct ideas of the relevant concepts. As the saying goes, the would-be rationalist should put up or shut up.

Let's look at one of them who did put up, or at least tried to. Frege's foundational views share at least some elements with traditional rationalism, and he made a valiant attempt to provide "proof, where proof is possible", of basic arithmetic principles. The details are well-known. He tried to found arithmetic and analysis on basic, logical principles concerning extensions (Frege [1884], [1893], [1903]). From the Fregean perspective, the question of this paper concerns the epistemic property had by those foundational principles for which no proof is possible.

Frege [1884, §2] tells us *why* it is that it is that mathematicians "prefer proof, where proof is possible":

> The aim of proof is, in fact, not merely to place the truth of the proposition beyond all doubt, but also to afford us insight into the dependence of truths upon one another. After we have convinced ourselves that a boulder is unmoveable, . . . there remains the further question, what is it that supports it so securely?

It is a Leibnizian perspective. To follow the metaphor, Frege believed that propositions have dependence relations to one another, perhaps along the lines of Bernard Bolzano's [1837] ground-consequence relation. The dependence relations are objective, in the sense that it is not a matter of how some person or other comes to believe that a given proposition, or even of how some person or other comes to know the proposition. Rather, it is a matter of what this truth rests upon.

Frege's account of the notions of analyticity and a priority are formulated in terms of the dependency relations:

> [T]hese distinctions between a priori and a posteriori, synthetic and analytic, concern, as I see it, not the content of the judgement but the justification for making the judgement. Where there is no justification, the possibility of drawing the distinctions vanishes. When . . . a proposition is called a posteriori or analytic in my sense, this is not a judgement about the conditions, psychological, physiological, and physical, which have made it possible to form the content of the proposition in our consciousness; nor is it a judgement about they way in which some other man has come . . . to believe it true; rather it is a judgement about the ultimate ground upon which rests the justification for holding it to be true.

> This means that the question is removed from the sphere of psychology, and assigned, if the truth concerned is a mathematical one, to the sphere of mathematics. The problem becomes . . . that of finding the proof of the proposition, and of following it up right back to the primitive truths. If, in carrying out this process, we come only on general logical laws and on definitions, then the truth is an analytic one . . . If, however, it is impossible to give the proof without making use of truths which are not of a general logical nature, but belong to the sphere of some general science, then the proposition is a synthetic one. For a truth to be a posteriori, it must be impossible to construct a proof of it without including an appeal to facts, i.e., to truths which cannot be proved and are not general . . . But if, on the contrary, its proof can be derived exclusively from general laws, which themselves neither need nor admit of proof, then the truth is a priori. (Frege [1884, §3])

It seems to me that despite the use of terms like "proof" and "justification", Frege's (and perhaps Bolzano's) relation of dependence is as much metaphysical as it is epistemic.[5] We have already seen that the relation has nothing to do with how people come to believe propositions. As Frege himself notes, it is also not a question of *whether* we know, for example, that 7+5=12. Skepticism and fictionalism aside, there is no serious question that we do know that. Nor is it a question of *how* we know that 7+5=12. That proposition was known long before the foundational work began. Moreover, for most of us, this knowledge need not, and in fact did not go via the proposed founding definitions. People just did the sum. Frege's dependency relationship thus seems to require a distinction between the state of knowing, or the state of being justified, and the *ultimate* or objective ground or justification of a proposition. The rationalist framework concerns the latter.

Let us call knowledge that is based on objective grounding relations among the known propositions *proper foundational knowledge*. For Frege, proper foundational knowledge is to be preferred. One advantage is that it uncovers the dependency relations among the propositions that one knows. We know *why* the boulder cannot be moved. Proper foundational knowledge is needed to determine whether a given proposition is analytic or a priori. Again, for Frege, is a matter of the proposition's metaphysical-cum-epistemic pedigree. One of the purposes of his logicism was to

---

[5]Thanks to Penelope Maddy and Michael Detlefsen here.

demonstrate that arithmetic and analysis are analytic. Robin Jeshion [2001, 940] calls this the "knowledge-of-sources rationale" for logicism.

Like Bolzano's ground-consequence relation, Frege's dependency relation is asymmetric: if proposition *A* depends on proposition *B*, then *B* does not depend on *A*. It follows that the relation is not reflexive: no propositions are grounded by themselves. Presumably, the relation is transitive. Can the dependency relation fail to be well-founded? Can there be infinitely descending dependency chains: A_0 depends on A_1 (plus perhaps other propositions), A_1 depends on A_2 (plus other propositions), A_2 depends on A_3, etc. Let us set aside the issue of whether it is coherent for someone to have proper foundational knowledge of each member of such a series. Frege's certainly never suggested such a thing; the proposed foundations are all finite.

It follows that some true propositions, so-called "primitive truths", are foundationally secure and not grounded on other propositions. His term is *selbstverstandlich*—self-standing.[6] With the possibility of non-well-founded dependencies aside, all properly foundational proofs must begin with self-standing propositions, as axioms. Certainly all finite properly foundational proofs must begin this way. Self-standing propositions require no proof and, indeed, no (non-trivial) proof them is possible. All other known propositions are based on self-standing propositions.

We come, then, to the present instance of the question of this paper. How are the self-standing, primitive truths known or, better, knowable? By definition, self-standing propositions are not (properly, foundationally) known on the basis of anything else. If we are to have proper foundational knowledge of the theorems—if we are to have proper foundational knowledge of anything—we must know the axioms. How?

---

[6]Some of the material that follows is drawn from the insightful studies in Detlefsen [1988], [1996, 56-70], Burge [1998] and Jeshion [2001], [2004] (see also Jeshion [2000]). Jeshion says that there is no proper English equivalent for *selbstverstandlich*, and she thus leaves the term untranslated.

For Frege, proper axioms have an epistemic property that he calls *einleuchten*, that of being self-evident. Jeshion [2001, 953] glosses the property as follows:

> (S-E) A proposition *p* is *self-evident* if and only if clearly grasping *p* is [a] sufficient and compelling basis for recognition of *p*'s truth.

As Tyler Burge [1998, 312] puts it, the relevant epistemic-metaphysical status is "something like *beyond a reasonable doubt by someone who fully understands the relevant propositions*". Let us call this notion *Fregean self-evidence*. The knowledge that comes via Fregean self-evidence cannot rely on any reasoning, or at least not on any deduction. The step from a proper grasp of the proposition to knowledge of it is immediate, and direct.

Jeshion sums up the aims of Frege's logicism as follows:

> **Euclidean rationale:** Frege thought that primitive truths of mathematics have two properties. (i) They are *selbstverstandlich*: foundationally secure, yet are not grounded on any other truth, and, as such, do not stand in need of proof. (ii) And they are self-evident: clearly grasping them is a sufficient and compelling basis for recognizing their truth. He also thought that the relations of epistemic justification in science mirrors the natural ordering of truths: in particular, what is self-evident is *selbstverstandlich*. Finding many propositions of arithmetic non-self-evident, Frege concluded that they stand in need of proof.

Some of this language is reminiscent of the Cartesian notion of clear and distinct ideas. By grasping the sense of a self-evident proposition, one thereby knows, without doing any more reasoning, that the proposition is true.[7]

With the traditional rationalist, Frege's metaphysical-cum-epistemic picture admires and puts a lot of weight on the power of the unaided human mind to discover truths about the world we find ourselves in. The picture strikes me as containing a large dose of pre-established harmony. According to the Fregean, there is a realm of propositions, which can be grasped by human beings. These propositions

---

[7]It does not follow that Fregean self-evident propositions are knowable on the basis of meaning alone (i.e., that they are analytic in the contemporary sense). Frege held that the axioms of Euclidean geometry are synthetic, with Kant. Yet, I presume, such axioms are self-standing and Fregean self-evident.

stand in objective dependency relations to each other, with self-standing propositions at the base. These self-standing propositions, or least the important foundational ones, happen to be self-evident, in the sense that grasping them is sufficient for knowing them, immediately. And these self-standing, Fregean self-evident propositions are sufficient for all mathematical knowledge.

Would that it were so. Even if there is a realm of propositions, distinct from and related to the sentences of natural languages, and even if these propositions enjoy objective dependency relations to each other, I see no reason why we should think they are structured in such a pleasing way, a way designed to facilitate proper foundational knowledge of the entire edifice of mathematics and, ultimately, science. It is natural to hope that the picture is correct. But do we have any reason to think it is?

A completed rationalist, Euclidean, Fregean program would go some way to defending this pleasing and harmonious metaphysical-cum-epistemic picture. Suppose, for example, that someone presented an axiomatization for arithmetic, real analysis, or set theory, in which the theorems obviously follow from the axioms and in which the axioms are obvious and, seemingly, not known on the basis of anything else (or at least knowable independently of any other knowledge). That would count as evidence that the axioms of the theory are Fregean self-evident, and that the deductions reflect objective dependency relations.

Of course, Fregean self-evidence is not obviousness, not a mere subject feeling of certainty. Too much reliance on obviousness would smack of the psychologism that Frege so vehemently opposed. There are obvious propositions, such as 7+5=12, that are not self-evident. To know that 7+5=12, one perhaps need not invoke any intuition, Kantian or otherwise, but one must *reason* one's way to this knowledge. You have to add it up, at least at first. More importantly, there are self-evident propositions that, at least at first, are not obvious. Even before he learned of Russell's paradox, Frege conceded that his Basic Law V is not obvious (of which more presently). In motivating the system, Frege [1964, 4] wrote:

> If we find everything in order, then we have accurate knowledge of the grounds upon which each individual theorem is based. A dispute can arise, so far as I can see, only with regard to my Basic Law concerning courses-of-values (V), which logicians perhaps have not yet expressly enunciated . . . Yet I hold that it is a law of pure logic.

To play its assigned foundational role, Basic Law V must be Fregean self-evident. Jeshion [2001, 960] argues that "Frege believed that what is self-evident does admit of doubt—even rational doubt". Tyler Burge [1998, 305] writes that Frege "expressed a sophisticated, modern awareness of the fact that what can seem obvious may turn out not even to be true . . . Frege was aware that principles that he put forward as axiomatic—even some that, unlike Axiom V, have endured as basic principles of logic—were not found to be obvious by his peers."

Nevertheless, the fact that a proposition is obvious is, presumably, some evidence that it is Fregean self-evident. One must start somewhere when setting up foundational programs, and at least some obvious propositions seem to be decent candidates for axioms. But as we have seen, obviousness is neither necessary nor sufficient for Fregean self-evidence. The inference from obvious to Fregean self-evidence is a fallible one, and there are Fregean self-evident propositions that, at least for a time, are not obvious. Burge [1998, 327] agrees that Frege "never says or implies that convictions about self-evidence are infallible. I think Frege believed that there is no infallible guarantee that one's commitments on logical or geometric truths are correct".

How is it possible for someone to understand a Fregean self-evident proposition and yet not know it? Given the above definition of self-evidence, the answer must be that it is possible to grasp a proposition in a less than perfectly clear manner. This is reminiscent of the rationalist notion of a confused grasp of concepts. Frege [1914, 216-7] considers a situation in which different mathematicians give different analyses of "number". Perhaps they are onto different notions, which happen to have the same name in natural language. Frege gives another take on the situation:

> [I]s not the explanation rather that we have really to do with the same science; that this man *does* attach the same sense to the word "number" as that man, only he doesn't manage to get hold of it

> properly? Perhaps the sense appears to both through such a haze that when they make to get hold of it, they miss it. One of them makes a grasp to the right perhaps and the other to the left, and so although they mean to get hold of the same thing, they fail to do so. How thick the fog must be for this to be possible!

So how *do* we get proper foundational knowledge of Fregean self-evident propositions. One possibility, I suppose, is to focus one's mind on the concepts and wait for clarity to come. This does not jibe with the "data". The mathematical giants who provided different "analyses" of, say, the real numbers, can hardly be faulted for failing to concentrate. Their minds are among the finest ever. How did some of them miss the concept, going to the left or the right in the haze?

Moreover, even when clarity *seems* to come, how do we *know* that it has. In general, how do we know, or when are we justified in believing, that the fog has lifted, that now we have a clear grasp of the concepts involved, and thus proper foundational knowledge? Why would Frege think that he, at long last, got it right? Jeshion [2004, 967] puts the issue well:

> As he was acutely aware of the possibility of errors resulting from conceptual understanding, Frege regarded reliance on obviousness as insufficient for identifying . . . primitive truths. As Frege noted, we are not given concepts 'in their pure form' [Frege [1884, vii]]. Our partial or incorrect understanding of concepts results in mistaken judgements. Such errors are not recognized 'from within' as, perhaps, are mistakes from inattention, sloppiness, or haste in judgement. And they are not remedied merely (!) by exercising control on one's thought, as, perhaps, are the others. The mistakes in question sometimes occur even when exercising tightest control on our intellection.

Apparently, Frege did not say in much detail how we properly come to know the starting points. How *do* we know that a given proposition is Fregean self-evident?

Burge [1998, 327] notes that "in arguing for his logic [Frege] made use of methods that were explicitly pragmatic and contextualist . . ." Here the dialectic takes a most interesting turn. Frege [1914, 205] praised the goal of organizing mathematical knowledge, presumably in a way that reflects the objective grounding relations:

> The essence of mathematics has to be defined from [a] kernel of truths, and until we have learnt what these primitive truths are, we cannot be clear about the nature of mathematics. If we assume that we have succeeded in discovering these primitive truths, and that mathematics has

> been developed from them, then it will appear as a system of truths that are connected to one another by logical inference
>
> Euclid had an inclination of this idea of a *system*; but he failed to realize it and it almost seems as if at the present time we were further from this goal than ever. We see mathematicians each pursuing his own work on some fragment of the subject, but these fragments do not fit together into a system; indeed, the idea of a system seems almost to have been lost. And yet the striving for a system is a justified one. We cannot long remain content with the fragmentation that prevails at present. Order can only be created by a system . . .
>
> . . . we must avoid such expressions as ‘a moment’s reflection shows that’ or ‘as we can easily see’. We must put the moment’s reflection into words so that we can see what inferences it consists of and what premises it makes use of. In mathematics we must never rest content with the fact that something is obvious or that we are convinced of something, but we must seek to obtain a clear insight into the network of inferences that support our conviction. Only in this way can we discover what the primitive truths are.

Based on passages like these, Jeshion [2004] argues that Frege’s methodology here is holistic. We come to know that a given proposition is self-standing and self-evident by examining its role in a carefully worked out scheme of knowledge. She has Frege

> advocating the sane view that what seems obvious may require proof and that obviousness needs supplementation by systematization. To identify a proposition as not needing proof [self-standing] . . . we need to systematize our knowledge and see whether the proposition can fulfil the role of an axiom within an ideal Euclidean system of mathematics. It does so by being fruitful, by enabling the derivation of all known mathematical knowledge and by affording means of generating more. It must also satisfy the traditional rationalist goals of ***surveyability, simplicity, economy, and unificatory power***. Jeshion [2004, 969]

See also Tappenden [1995].

As noted, even before he heard from Russell, Frege conceded that Basic Law V is not obvious, and that one can rationally doubt that it is Fregean self-evident (or, indeed, even doubt its truth). In [1898, 182], he wrote that:

> by what right does such a transformation take place, in which concepts correspond to extensions of concepts . . . An actual proof can scarcely be furnished. We will have to assume an unprovable law here. Of course it isn’t as self-evident as one would wish for a law of logic. And if it was possible for there to be doubts previously, these doubts have been reinforced by the shock the law has sustained from Russell’s paradox.

Jeshion provides two “charitable” interpretations of Frege’s pre-Russell inclusion of Basic Law V as an axiom. One is that Frege envisioned the possibility that, one day, the principle might be proved from other, even more basic propositions. That is, we might be able to analyze the concept of “extension” into more primitive terms, and prove Basic Law V from self-standing, self-evident propositions about those terms. At the time, he did not know how to prove the principle and so, temporarily, he took it as an axiom.

With regard to the theme of this paper, Jeshion’s other “charitable” interpretation is more intriguing:

> Frege could have held that the coincidence in sense . . . would, in time, come to seem as obvious as the other axioms. Basic Law V seemed true even thought it came up short on the obviousness needed for thinking it self-evident . . . But, as Frege recognized, that could change . . . Frege . . . recognized that [the terms] were not fully understood. By proving theorems with the new basic concepts, understanding may well alter enough so that Basic Law V would come to possess the obviousness required for thinking it self-evident. (Jeshion [2004, 290-291])

Jeshion makes an analogy with the ϵ-δ definitions of such notions as convergence, continuity, and differentiability. They were not obvious at first, but in time, it became clear that they are correct, to the extent that the definitions are now obvious. Surely, this clarity emerged through extensive *use* of the definitions in proofs.

This is exactly what was said, in the previous section, when we considered self-evidence as a species of obviousness, as type of psychological certainty. The axioms of set theory, and of most other (assertory) branches of mathematics, do indeed “force themselves on us as true”, as Gödel put it, but this forcing does not occur on a first or second reading of them. The axioms only become certain after immersion in the practice. So do the theorems.

Burge [1998, 328] also sounds the holistic theme, perhaps less explicitly:

> In *Basic Laws* [1893] we find Frege recommending to those who are sceptical of his logical system that they get to know it from the inside. He thinks that familiarity with the proofs themselves will engender more confidence in his basic principles . . . In the Introduction to *Basic Laws*, Frege repeatedly appeals to advantages, to simplicity, and to the power of his axioms in

> producing proofs of widely recognized mathematical principles, as recommendations of his logical axioms.

A bit later, Burge adds that "whatever role self-evidence plays in [Frege's] epistemology seems to be qualified by pragmatic considerations that result from reasoning within and about his system of proofs over time" (p. 335).

On the surface, at least, there is some tension in this picture. Burge [1998, 315] writes that "self-evidence must partly be understood in terms of recognizability as true *independently of recognition of other truths*. Sufficient evidence to make believing them rational *is carried in these individual truth themselves*" (my emphasis). This is in line with the above definition of Fregean self-evidence, from Jeshion [2004]. But, according to these interpretations, we sometimes get into the position of knowing a proposition independent of other truths only when we see the role of the proposition in a systematization of those other truths. Well, which is it? Are the axioms known individually, in themselves, or are they known holistically, in terms of their role in a successful system?

In sum, Frege elaborated a complex metaphysical-cum-epistemic system, according to which some propositions are objectively grounded on others. When faced with the question of how one comes to know the axioms, the propositions that lie at the base of the edifice, we are told that they are self-evident: one who clearly and fully understands them realizes that they are true, without doing any reasoning. But understanding is sometimes confused, or otherwise less than sufficiently clear. How does one get the proper understanding and thus knowledge of the axioms? How does one get a fully clear grasp of the concepts. At this point, according to Jeshion and Burge, Frege goes systematic. We get a fully clear grasp of concepts by working with them, by immersing oneself in the practice of deducing the consequences of these axioms, and seeing how everything fits together.

Burge resolves the tension here with an interpretative proposal that in some cases, it is only an *ideal mind* that can know a self-evident proposition independently of any other. We fallible, mortal humans invoke systematic considerations to obtain evidence about what an ideal mind can know without reasoning, independent of other propositions. Concerning the pre-Russell status of Basic Law V, Burge wrote: "Since [Frege] thought it was an axiom, he must have, at least sometimes, thought that *it* was certain, but because of insufficient analysis or incomplete understanding, *he* was not". Frege "maintained the views, which he several times expresses in the pre-paradox period, that the basic principles that he proposed are genuine axioms and that axioms are self-evident" (Burge [1998, 337-338]).

This distinction is reminiscent of Leibniz's proposal that God knows every proposition via a clear grasp of the concepts it contains, but us mortals have to know things in a more indirect manner. For the Fregean rationalist, it seems, the proper epistemic basis of proper axioms is self-evidence, as above. Understanding them is indeed sufficient for knowing them—ideally. But in practice, humans are not so fortunately situated. We have to rely on holistic or systematic considerations before we are assured of the axioms.

Let *p* be a proper axiom. Perhaps the idea is that even humans know *p* independent of any other propositions, on the basis of a clear understanding of *p*. But sometimes we need holistic or systematic considerations to assure ourselves *that p* is self-evident, and thus knowable independently of other propositions. That is, if *p* is Fregean self-evident, then our knowledge of *p* is indeed unmediated by anything else, but sometimes we only know that *p* is self-evident, and unmediated by anything else, in a mediated way, by systematic considerations. There is a difference between how we know *p* and how we know how we know *p*.

As we saw, Jeshion [2004, 209] adds that holistic coherence is not sufficient for knowing that a given proposition *p* is Fregean self-evident: *p* "must also satisfy the traditional rationalist goals of *surveyability, simplicity, economy, and unificatory power*". Frankly, I see no real role here for surveyability here, especially if we take Burge's line on ideal knowers.[8] It would add to the pre-established harmony if, in addition to being self-standing and Fregean self-evident, the primitive truths of mathematics were surveyable. Why think this? Why think that the propositions that lie at the base of the objective dependency relation happen to be surveyable by minds like ours? What is special about the powers of memory and concentration of human mathematicians? Suppose that a given proposition is too complicated for us to understand it in one glance. How can that alone disqualify its foundational status?

The other three criteria that Jeshion lists—simplicity, economy, and unificatory power—are not particularly rationalist. Just about everyone lists these as theoretical virtues. For the staunch empiricist W. V. O. Quine, for example, simplicity, economy, and unificatory power are scientific criteria for the web of belief. From the rationalist perspective, some of these criteria smack of even more pre-established harmony. Why think that the propositions that lie at the base of the dependency relation are simple? Indeed, why think that the mathematical universe itself is simple, and thus that we humans can easily grasp its basic principles. Why think that the objective base is, in fact economical? Why think it is unified? Was it built to make life easy for creatures with minds like ours?

It seems to me that once we go holistic, either at the level of knowledge (for flesh and blood humans) or at the level how we know a proposition's status, we compromise the goals and indeed the spirit of rationalism. One can wonder what is left of the properly rationalist or Euclidean elements in the Fregean system. If, on Jeshion's and Burge's readings of Frege, we are to go holistic at the crucial place, then why not wax holistic from the start? Why don't we claim that we know the whole system—axioms and theorems together—in terms of the role it plays in our overall conceptual scheme, in our form of life, or in the web of belief? The burden on a would-be Euclidean who follows this route is to show that the

---

[8]I presume that the proper notion is "surveyable by a human". I do not know what it would mean for a proposition to be ideally surveyable. I presume that all (finitary) propositions are ideally surveyable.

elements of objective dependence and Fregean self-evidence have some explanatory role to play. What light do they shed on the nature of mathematics?

Since at least Gödel, we have learned to fly without a safety net. To change the metaphor, if the edifice of mathematics is somehow grounded in bedrock, it is not indubitably so. Why think that there are objective grounding relations at all? What role do they play in philosophy of mathematics?

This is not the place, and I am not the author, to explore every possible Euclidean or rationalist account of mathematics. The burden on the Euclidean is to give an account of how we know the axioms, the propositions that lie at the basis of the system. Presumably, they are self-evident. What does this mean? And how do we know *that* the axioms are self-evident? Is that self-evident as well?

Before moving on, let us briefly consider the so-called Scottish school of neo-logicism (e.g., Hale and Wright [2001]). Frege's *Grundlagen* [1884] contains the essentials of a derivation of the Peano postulates (and thus the rest of elementary arithmetic) from a premise that was once dubbed $N^{=}$ in Wright [1983] and is now called *Hume's principle*:[9]

$$(Nx{:}Fx = Nx{:}Gx) \equiv (F \approx G),$$

where $F \approx G$ is an abbreviation of the second-order statement that there is a one-to-one relation mapping the F's onto the G's.

The derivation in question, now called *Frege's theorem*, reveals that Hume's principle entails that there are infinitely many natural numbers, that induction holds on them, etc. It is generally agreed that this is a powerful mathematical theorem. The neo-logicist proposes to take this success story as an Euclidean foundation for arithmetic. Building on that, the program is to found other branches of mathematics on other principles in the form

[9] A more apt name would be "Cantor's principle", giving credit where credit is due, but "Hume's principle" has become standard.

$$\text{(ABS)} \qquad \forall a \forall b (\Sigma(a)=\Sigma(b) \equiv E(a,b)),$$

where *a* and *b* are variables of a given type (typically individual objects or properties), Σ is a higher-order operator, denoting a function from items of the given type to objects in the range of the first-order variables, and *E* is an equivalence relation over items of the given type. Propositions in this form are called *abstraction principles*.

Let us stick to Hume's principle. This is the proposed foundational axiom for arithmetic. So, to ask the question of this paper, how is it known? Is it self-evident in some sense? Hale and Wright [2000] propose that Hume's principle is akin to an implicit definition, and is true by stipulation. In laying down Hume's principle, we are merely saying what we mean by the "number of" operator. The program thus has something in common with the Hilbert-style, algebraic implicit definitions considered in §1 above. Unlike the algebraist, however, the neo-logicist *asserts* Hume's principle, taking it to be a truth about cardinal numbers and the number-of operator. He claims to have a priori knowledge of Hume's principle, via the stipulation. Frege's theorem thus reveals that the basic propositions of elementary arithmetic are known a priori, on the basis of their derivation from the foundation.

Frege's Basic Law V is another abstraction principle:

$$(\mathrm{E}x{:}Fx = \mathrm{E}x{:}Gx) \equiv \forall x(Fx \equiv Gx).$$

Unlike Hume's principle, Basic Law V is inconsistent, and thus cannot serve as a foundation for anything. We certainly don't know *it* a priori, since we do not know it at all. The formal resemblance between Hume's principle and Basic Law V has exploited in a so-called "bad company" objection to the Scottish neo-logicist program (see Boolos [1997]).

Like Jeshion's and Burge's Frege, the neo-logicist notes that the foundational method is fallible. There is some risk in involved in stipulating an abstraction principle; the enterprise might go wrong. It seems to work OK with Hume's principle and arithmetic, but it would fail if someone tried to stipulate Basic Law V and found a theory of extensions (and a set theory) on that.

So which are the allowable stipulations? How do we tell? Hale and Wright have risen to this challenge and delimited properties that a successful abstraction principle must have. Consistency is an obvious necessary condition, and they have suggested some others, such as various conservativeness requirements.

In response to the question of the present paper, then, the neo-logicist declares that the proper foundational axioms are known a priori, by stipulation. But to fill that role, certain conditions must be met. The abstraction principle must be consistent, conservative, etc. It is surely too much to demand that the neo-logicist *prove* that the proposed axiom is consistent, conservative, etc. Even putting the second incompleteness theorem aside, it would undermine the Euclidean enterprise if we had to *prove* something before we could make the foundational stipulation. How would that purported proof go? What are its premises or axioms?

Surely the neo-logicist owes us some account of how we know—or at least how we are assured or why we are justified in thinking—that a given stipulation is correct, that the proposed conditions are met. This is an ongoing research interest of the program, and so I cannot be definitive here. One natural response is to follow Jeshion's and Burge's Frege and wax holistic. We become assured that a given abstraction, such as Hume's principle, is consistent, conservative, etc., and thus known a priori by stipulation, via the role of the abstraction in a systematization of mathematics, or a branch thereof (see Rayo [2003]).

Once again, the goal of the program is not to give mathematics a safety net. We can't have that in any case, thanks to Gödel. One goal shared by logicism and neo-logicism is to reveal the *sources* of mathematical knowledge. It is not a matter of wondering *whether* we know something, but of saying something about *how* we know. Why is the boulder immovable? But, as with Jeshion's and Burge's Frege, if we go holistic when it comes to how we know that the axioms are axioms, and thus foundationally secure, then why not go holistic from the start? Why not take the whole system, basic arithmetic, analysis, set theory, whatever, to be known via its role in our intellectual lives. What explanatory work do the Fregean and neo-logicist elements play in the Euclidean enterprise? Again, I do not claim to be definitive here, just to have raised an interesting and important question for the Euclidean program.

4 Zermelo: self-evidence as unconscious use. Ernst Zermelo's celebrated axiomatization of set theory appeared in his [1908], which also contains a new proof of the well-ordering theorem. The bulk of that article is a section (§2) entitled "Discussion of the objections to the earlier proof" of that theorem (Zermelo [1904]). In the first sub-section, "Objections to the principle of choice", Zermelo claims that choice is "self-evident". For purposes of this discussion, let us focus on the following version of the axiom of choice:

(AC) If x is a set of non-empty sets, then there is a function f such that for each $y \in x$, $fy \in y$.

It is interesting that Zermelo maintained that choice is self-evident even after many, probably most, of the leading mathematicians of the day—notably the French analysts Baire, Borel, and Lebesgue—balked at it (see Moore [1982]). Some thought it false, while others expressed doubts, a healthy agnosticism perhaps.

By "self-evident", Zermelo did not mean "obvious". One can, of course, hold that a given proposition p is true despite that fact that a large number of respectable thinkers are not sure, express doubts, or even think it false. But under these circumstances, one should not maintain that p is obvious. I presume that Zermelo respected his opponents, and was not claiming that they are so dense as to fail to grasp what should be obvious. The paper in question is a serious mathematical and philosophical enterprise, not an exercise in name-calling.

It is sometimes thought that the issue turns on different philosophical conceptions of mathematics and mathematical objects (see Maddy [1990, 121-123], Lavine [1994, 111-115]). If someone holds a constructivist philosophy, maintaining that one must show how to define or construct an object before claiming that it exists, then, presumably, she will doubt choice. The “axiom” does not provide a construction or definition of a choice function; it only asserts that such a function exists. The non-constructivist, or “combinatorial”, view is that sets are arbitrary collections of objects, and that functions are arbitrary correspondences. Sets and functions exist whether or not they can be defined or constructed. Presumably, choice *is* obvious on the combinatorial view.

Historically, there is some truth to this diagnosis. The French analysts did have constructivist leanings, and Zermelo and some of his supporters, such as Hadamard, did not, separating out questions of existence and definability. Moreover, one can see a trend in the early decades of the twentieth century away from a constructivist perspective on mathematics, along with a trend toward acceptance of the axiom of choice.

Nevertheless, this is not the entire explanation of the situation on the ground in the early twentieth century, nor is it particularly enlightening in understanding Zermelo’s claim of self-evidence. Once again, Zermelo never claimed that choice is *obvious*, as it supposedly is on the combinatorial view. At the time, issues of constructivity and definability were not as clearly formulated as they are today, and it is not straightforward to characterize the players as constructivist and combinatorialist, nor is it clear that such a lineup would correspond exactly with stances on choice.

Moreover, at least some versions of choice are true—and obviously true—on a thoroughly constructivist reading. Consider the antecedent of (AC): x is a non-empty set of sets. On a constructive reading, this entails that there is a uniform method for constructing a member of each member of x. For the constructivist, this is what it is for x to *be* a set of non-empty sets. Intuitively, this method will produce the choice function, at least if identity is decidable.[10]

[10]The issue of choice from a constructivist perspective is much more subtle than this. In general, if identity is not decidable, then choice is not true. For example, choice fails for the constructive theory

From the other perspective, I would suggest that choice is not quite obvious from a combinatorial point of view. Or at least a reasonable doubt can be raised. For one thing, the axiom has some rather counterintuitive consequences, such as the Tarski-Banach theorem. Moreover, typical arguments in favor of (the obviousness of) choice beg the question. Consider, for example, the Cartesian product *p* of a given (infinite) set *x*. Surely, it is obvious that *p* is not empty. But the statement that *p* is non-empty just is the axiom of choice. And, in general, we know that infinite sets do not share all the properties of finite sets. Or suppose that *x* is a (infinite) set of pairwise disjoint non-empty sets. Let *q* be the union of *x* and let *Pq* be the powerset of *q*. Since *Pq* is the set of *all* subsets of *p*, then surely—it is obvious that—one of those subsets consists of exactly one member of each member of *x*. But, again, this just begs the question.

In any case, Zermelo's defense of (AC) does not go via this route. He never claims that any careful thinker will be convinced of its truth just by thinking about what it says. To use the above terminology, Zermelo does not claim that (AC) is Fregean self-evident. So what did Zermelo mean by "self-evident"? In his [1904] proof of the well-ordering theorem, he claimed that the axiom of choice is a "logical principle", and that it "cannot . . . be reduced to . . . still simpler" logical principles. That is, the principle of choice cannot be proved, and so it is what Frege called *selbstverstandlich*, self-standing. But, Zermelo claimed, "it is applied without hesitation everywhere in mathematical deduction". Zermelo cites an example: "the validity of the proposition that the number of parts into which a set decomposes is less than or equal to the number of all of its elements cannot be proved except by associating with each of the parts in question one of its elements".

In the [1908] axiomatization, Zermelo repeats his 1904 assertion that he cannot prove the axiom of choice, and adds that he "therefore cannot compel anyone to accept it apodictically" (§2.a). But he

---

of real numbers, and it fails in set theory, if extensionality is assumed.

quickly points out that "in mathematics ***unprovability*** . . . is in no way equivalent to ***nonvalidity***, since, after all not everything can be proved, but every proof in turn presupposes unproved principles". This broaches the theme of the present article. How do we know these unproved principles? And how do we know which principles those are?

Zermelo goes on to state that "in order to reject such a fundamental principle, one would have to ascertain that in some particular case it did not hold or to derive contradictory consequences from it; but none of my opponents has made any attempt to do this". Although this seems to shift the burden of proof (so to speak) to his opponents, Zermelo does go on to justify his inclusion of a version of (AC) as an axiom. He observes that even "Peano's *Formulaire* [1897] . . . rests upon quite a number of unprovable principles". Zermelo claims that Peano arrives "at his own fundamental principles . . . by analyzing the modes of inference that in the course of history have come to be recognized as valid and by pointing out that the principles are intuitively evident and necessary for science". Here is our phrase again. What is it to be "intuitively evident"?

Zermelo then claims that these considerations can "be urged equally well in favor of the disputed principle", namely the axiom of choice. He supports the claim that choice is "necessary for science" by providing a list of seven theorems that seem to rely on choice (one of which was the example used in the [1904] note).

This sounds a holistic theme, which we have encountered already, several times. But what does that have to do with "self-evidence" or "intuitive evidence"? Zermelo continues:

> That this axiom, even though it was never formulated in textbook style, has frequently been used, and successfully at that, in the most diverse fields of mathematics, especially in set theory, by Dedekind, Cantor. F. Bernstein, Schoenflies, J. König, and others is an undisputable fact, which is only corroborated by the opposition that, at one time or another, some logical purists directed against it. Such an extensive use of a principle can be explained only by its *self-evidence* . . . No matter if this self-evidence is to a certain degree subjective—it is surely a necessary source of mathematical principles, even if it is not a tool of mathematical proofs . . .

Let us call a proposition ***Zermelo self-evident*** if it is applied instinctively, without explicitly citing it, in a wide variety of instances. Besides being vague, this notion is clearly relative. I should perhaps say that a proposition *p* is Zermelo self-evident *for a mathematician S at time period t* if *S* applies *p* instinctively during *t*, without explicitly citing it, in a wide variety of instances. A proposition can become Zermelo self-evident as practices evolve, and a proposition can lose its status as Zermelo self-evident if one stops applying it instinctively. In what follows, I will usually drop the relativity, leaving it implicit. In the cases of interest, the "mathematician" is the mathematical community at large, in a certain historical context.

Presumably, logical principles are Zermelo self-evident. Except in logic classes, we apply or-elimination, excluded middle, non-contradiction, and the like all of the time, without citing them and often without being aware that we are relying on the principle in question. Zermelo claims, then, that the axiom of choice is Zermelo self-evident, and he supports this with a number of examples.

Cantor's famous 1899 letter to Dedekind, reprinted in van Heijenoort [1967, 113-117], contains an argument for a version of the well-ordering theorem. Cantor shows how to embed ordinals into any given "definite multiplicity" via a transfinite series of choices. Zermelo, who served as editor of Cantor's collected works [1932], added a note criticizing the derivation:

> . . . the intuition of time is applied here to a process that goes beyond all intuition, and a fictitious entity is posited of which it is assumed that it could make *successive* arbitrary choices and thereby define a subset . . . that . . . is precisely *not* definable. Only through the use of the "axiom of choice", which postulates the possibility of a *simultaneous* choice and which Cantor uses unconsciously and instinctively everywhere but does not formulate explicitly anywhere, could [the subset] be defined . . . It is precisely doubts of this kind that impelled the editor a few years later to base his own proof of the well-ordering theorem [Zermelo 1904] purely upon the axiom of choice . . .

(p. 17) Here we see Zermelo claiming that, for Cantor at least, (AC) is Zermelo self-evident.

To be sure, what I call Zermelo self-evidence is not exactly what Zermelo meant by "self-evidence". He said that the fact that choice is applied instinctively and widely is *explained* by its self-evidence (Zermelo [1908, §2a]). That is, Zermelo argues that (AC) is self-evident as an inference to the best explanation of the fact that (AC) is what I call Zermelo self-evident.

Zermelo's writings contain enough hints and explicit phrases to attribute a platonic ontology to him. In the terminology of §1 above, for Zermelo, set theory is assertory, and its subject matter is an objectively existing, eternal, abstract realm of sets. He is not a Hilbert-style algebraist. In an intriguing and insightful article, G. Gregory Taylor [1993] suggests, at least tentatively, that Zermelo also held a Platonic epistemology, of sorts: mathematicians have some access to the platonic realm of sets. In discussing impredicativity, Zermelo [1908, §2b] wrote that "the point of view maintained here [is] that we are dealing with a productive science resting ultimately upon intuition". But, as with rationalist foundations, the access, or "intuition", is sometimes foggy. Taylor writes: :

> Zermelo's historical method for establishing that axioms are intuitively evident suggests that mathematicians must have access to some platonic domain of sets in its entirety. This access will not be direct or immediate . . . In this respect, Zermelo is probably not unusual . . . More novel is Zermelo's idea that evidence for this access—whatever its nature (and there is little point in speculating on what Zermelo takes the nature of this access to be)—is gathered *empirically* by examination of the work of practicing mathematicians. (Taylor [1993, 557])

The idea, then, is that the instinctive, unconscious use of choice counts as quasi-empirical, a posteriori evidence that mathematicians do "perceive" that it is true. At some perhaps unconscious level, it is indeed intuitively evident. We know this because it is Zermelo self-evident.

I propose to leave behind the proposal for reading Zermelo as endorsing a Platonic epistemology. Present interest is also somewhat independent of the ontological issues as well. If someone asserts the axioms of set theory, then, of course, she holds that these axioms and thus the theorems are true. But we need say little more about set theory's distinctive subject matter—if it has one. What matters, at present, is whether and how we know the axioms, and thus the theorems. I make the modest proposal that the notion of Zermelo self-evidence is interesting and important, and deserves some consideration.

Zermelo's insight that in the early twentieth century, the axiom of choice is Zermelo self-evident was quite correct, and in the end, decisive. He ended the [1908] paper with the "hope that in time all of [the] resistance can be overcome through adequate clarification". This is something Frege might have said about Basic Law V, as well. But here it happened. To simplify a complex story, the mathematical community came to realize that the proofs of many major theorems rely on choice principles. Ironically, this included the work of the French analysts, Borel, Borel and Lebesgue—the major opponents of the axiom of choice. Choice was indeed Zermelo self-evident, even for them. This led to an intense effort to prove the theorems without relying on choice, efforts which often failed. In many of the cases, it was seen—proved—that the use of choice was necessary, often that the theorem in question implies a choice principle. Over the ensuing decades, the mathematical community at large, with fewer and fewer exceptions, came to accept the axiom of choice—despite its untoward consequences.[11] What was done unconsciously and instinctively is now down explicitly and consciously, without blinking.

Notice that a proposition can be Zermelo self-evident without being "evident" in any intuitive sense. It need not be obvious. Moreover, once a Zermelo self-evident proposition is made explicit, a thinker can balk at it, or even reject it outright, as happened with the axiom of choice. As we have seen, this can happen even if it is *shown* or *made explicit that* the proposition in question is Zermelo self-evident.

For what it is worth, a similar phenomenon occurs outside of mathematics. In ethical and political considerations, for example, the doctrine of double-effect is, roughly, that an agent is not culpable (or less culpable) for foreseeable but unintended consequences of his acts. Once the doctrine is made explicit, it is debatable, to say the least. However, it is my experience, at least, that many people apply this principle instinctively and unconsciously in dispassionate discussions. They speak of

[11] Moore [1982] is a detailed account of the history. It makes for fascinating reading.

collateral damage, and the like. When the "principle" is made explicit, they sometimes withdraw the argument, or else go on to defend the doctrine of double-effect.

Getting back to mathematics, notice that a false or even contradictory proposition can be Zermelo self-evident. Arguably, Frege's Basic Law V, or some other version of an unrestricted comprehension principle, is or was Zermelo self-evident. In so-called naive set theory, it is applied all the time. So is the principle that every well-ordering has an order type and, at one point in history, the proposition that a subset of a set is smaller than the set, and the proposition that a continuous function is differentiable almost everywhere (or at least somewhere).

Clearly, then, the method of choosing axioms that are Zermelo self-evident is defeasible. But no one ever said it isn't; we are not looking for a safety net. Zermelo [1908, §2a] himself noted that "there are no infallible authorities in mathematics", echoing a theme we encountered throughout this article. As we saw, Zermelo conceded that he "cannot compel anyone to accept [the axiom of choice] apodictically" (§2a). Yet, when developing axiomatic systems, one has to start somewhere, and the fact that a proposition is Zermelo self-evident is sufficient reason to include it, at least provisionally. Or so it seems to me. As Zermelo [1908, §2a] wrote:

> . . . so long as . . . the principle of choice cannot be definitively refuted, no one has the right to prevent the representatives of productive science from continuing to use this "hypothesis"—as one may call it for all I care—and developing its consequences to the greatest extent, especially since any possible contradiction inherent in a given point of view can be discovered only in that way . . . Banishing fundamental facts or problems from science merely because they cannot be dealt with by means of certain prescribed principles would be like forbidding the further extension of the theory of parallels in geometry because the axiom upon which this theory rests has been shown to be unprovable.

Axioms, whether they are obvious, intuitively evident, or Zermelo self-evident, must pay their dues by playing a role in a systematization of an established and successful practice. The above passage concludes: ". . . principles must be judged from the point of view of science, and not science from the point of view of principles fixed once and for all". So we are in a situation much like the one we found above with Jeshion's and Burge's Frege, and the concomitant notion of Fregean self-evidence.

To sum up, Zermelo employs the traditional Euclidean, rationalist vocabulary of "justification" and speaks of what is and what is not "provable". As Taylor [1998, 598] notes, "there seems little doubt that his intention is to . . . ground mathematics in an epistemic sense". I presume this invokes something like Fregean objective grounding relations. Taylor continues,

> . . . when it comes time to defend the axioms, Zermelo adopts a problematic stance . . . No doubt Zermelo regards some of his axioms as straightforwardly self-evident. It is clear, however, that he does not regard them all that way, since what he emphasizes is indispensability. We can determine objectively through the examination of mathematical argumentation presented in written texts that a proposition has often been appealed to in the past . . . The larger question from the point of view of foundations is this: How are the axioms to ground mathematics if our best evidence for them is that very mathematics? What seems to emerge is a conception of foundations that is not Cartesian at all . . .

Just as we saw with Fregean self-evidence, we can wonder what is left of the Euclidean, rationalist, foundationalist image. By definition, we cannot provide non-trivial proofs of axioms, at least in the system in which they are axioms. But, presumably, we do know these axioms. If we have to turn holistic when asked *how* we know them, then what is the point of the rationalist framework? Why not just say that we know the axioms and the theorems alike via their role in systematizing our mathematical theories?

I do not claim to have canvassed every possible notion of self-evidence in play at the foundation of an assertory, Euclidean, rationalist system that underlies a sophisticated mathematical theory. Obviousness alone is too psychological and too ephemeral to play the foundational role. Besides, many of the actual axioms are not obvious, independent of immersion in the practice. We saw that Zermelo self-evidence is not an infallible guarantee of truth, and we have no general infallible route to knowing that a given proposition is Frege self-evident. If these notions must be supplemented by holistic considerations, then one can wonder what the goals of the rationalist, Euclidean enterprise are, or what the epistemic weight of self-evidence is.

5 Final word. The "Declaration of Independence of the Thirteen Colonies" that would later constitute the United States of America reads, "We hold these truths to be self-evident: that all men are created equal, that they are endowed by their Creator with certain unalienable Rights, that among these are Life, Liberty and the pursuit of Happiness". Let us leave aside the part about the Creator. I have no doubt that all people should indeed be treated as equal in any civilized community, and that the human rights listed are inviolable. But I confess to having very little idea of what the author of this venerable document meant by "self-evident". The propositions are not Zermelo self-evident, applied instinctively, although it would be a much better world if they were. Perhaps the truths are Frege self-evident. Perhaps a sufficient grasp of the propositions, and in particular a clear and distinct knowledge of the word "men" (or "people") is sufficient for one to know, without doing any reasoning, that the propositions are true. Would that it were so. Like the axiom of choice, the truths in question are indeed obvious now, at least to many of us, but they were hardly obvious then. If they were obvious, the framers would not even have to state them, much less call them self-evident.