Review: "Foundations Without Foundationalism, A Case for Second-Order Logic" S. Shapiro, Clarendon Press, Oxford, 1991

Nino B. Cocchiarella

Foundations Without Foundationalism, A Case for Second-Order Logic (Oxford University Press, Oxford, 1991), by Stewart Shapiro, is an excellent book, covering of all of the main results in second-order logic and its applications to mathematical theories. Its main theme is that first-order logic does not adequately "codify the descriptive and deductive components of actual mathematical practice", and that "first-order languages and semantics are also inadequate models of mathematics" (43). Second-order logic (under its "standard" semantics), Shapiro maintains, "provides better models of important aspects of mathematics, both now and in recent history, than first-order logic does" (v); and in that regard it is second-order, and not only first-order, logic that "has an important role to play in foundational studies" (ibid.). Indeed, the restriction of logic to first-order logic (without Skolem relativism) in such studies is "the main target of this book" (196).

The book is divided into three parts, with Part I containing a discussion of the philosophy of logic and the role of logic in foundational studies. In Part II, a technical development of second-order logic is given in which it is argued that "higher-order notions are well-suited for modelling important aspects of mathematics" (x). Additional philosophical issues relevant to the acceptance of second-order logic are given in Part III, as well as a coverage of some of the history of these issues in the early part of this century.

1 Full Logics

By a *full logic*, Shapiro means "a [formal] language, together with a deductive system and a semantics" (3) – where the language is taken as a model of a fragment of ordinary natural language considered as a natural language of mathematics. Given a set K of nonlogical symbols (individual, function, and predicate constants), the main full logics considered here are $L1K{=}$, first-order logic with identity, and $L1K$, first-order logic without identity, respectively – as based on the symbols in K (with $\rightarrow$ for the material conditional, $\neg$ for classical negation, $\forall$ for universal quantification, and $\exists$ and other truth-functional connectives defined in the usual way). $L2K-$ is $L1K=$ extended to contain predicate and function variables of different degrees, but no quantifiers binding these; and $L2K$ is $L1K$ similarly extended, but with quantifiers binding predicate and function variables as well as individual variables. (Identity is primitive in $L2K-$, but defined in $L2K$ as indiscernibility.) "Standard" deductive systems, $D1=$ and $D1$ are described for the first-order logics $L1K=$ and $L1K$, and deductive systems $D2-$ and $D2$ extending $D1=$ and $D1$ are described for $L2K-$ and $L2K$, respectively. Having no bound predicate and function variables, $D2-$ has no comprehension principle, but it does have a substitution rule (to infer $\Phi[R/\Psi(\langle x\rangle_n)]$ from Φ).

The distinctive new axiom schemes for $D2$ are (1) the full impredicative comprehension principle, (CP), for concepts (including relational concepts),

$$\exists X^n \forall \langle x\rangle_n (X^n \langle x\rangle_n \equiv \Phi\langle x\rangle_n),$$

which (for X^n not free in Φ) says, in effect, that every formula of $L2K$ determines an n-ary concept (for each positive integer n), (2) a similar principle for functions (taken primitively, but which could be eliminated in favor many-one relational concepts), and (3) an axiom of choice.

No justification or explanation otherwise is given as to why an axiom of choice is considered a logical truth of second-order logic – though its validation is automatic in the author's metalanguage ZFC, Zermelo-Fraenkel set theory, ZF, with an axiom of choice, C. The result of dropping the axiom of choice from $D2$, called $D2^*$, is what some writers (including this reviewer) call "standard" (impredicative) second-order logic in the syntactical, as opposed to the set-theoretic, semantical sense. (It is also usually assumed that primitive functions have been eliminated in $D2^*$ in favor of many-one relations, thereby reducing the system to a formal logic of predication in which functionality is not represented as a separate category.) It is this system that is equivalent to the second-order logic of Frege's *Begriffsschrift*, and it is in this sense that we can say that Frege was the first to develop not only "standard" first-order logic (as a proper part of his *Begriffschrift* system), but also "standard" second-order logic as well. Shapiro does not discuss this system, and no explanation is given of why Frege thought that the system embodied all of the "fundamental laws of logic" (other than his law for the extensions of concepts). The main reason for ignoring the system, apparently, is that it falls far short of determining all of the logical truths (or all of the valid arguments) of second-order logic under its so-called "standard" set-theoretical semantics. (But then so does $D2$ as well.)

The semantics of the full logics $L1K$ and $L1K=$ is based on set-theoretic structures called models. In particular, a model for $L1K$ and $L1K=$ is a 2-tuple $\langle d, I\rangle$, where d (called the *domain,* or *universe, of discourse* of the model) is a nonempty set, and I is a function (called an "*interpretation*") assigning entities "constructed from d" of the appropriate types to the nonlogical symbols in K. The satisfaction in such a model of a first-order formula by an assignment of values drawn from d to the first-order variables is understood in the usual way, and the logical truth (or semantic validity) of a formula is its satisfaction in every model by all assignments in those models. The validity of an argument is defined similarly.

The same set-theoretic models are taken as the "standard" models of $L2K-$ and $L2K$, only now the assignments of values to variables are extended to predicate and function variables as well – namely, by assigning subsets of d^n (the set of n-tuples drawn from d) to n-place predicate variables, and functions from d^n to d to each n-place function variable. What makes this semantics "standard" for $L2K$ and $L2K-$ is that n-place predicate variables range over the entire powerset of d^n (as determined by Cantor's theorem) and n-place function variables range over all of the functions in d^{d^n}. Formulas are defined as "*standardly valid*", or as "*standard logical truths*", if satisfied in every model by all such assignments in those models, and the "standard validity" of arguments is defined similarly. "In standard semantics", Shapiro writes, "by fixing a domain one thereby fixes the range of both the first-order variables and the second-order variables" (73) – as determined in the set-theoretic metalanguage by the Cantorian notion of powerset.

In the Henkin, or so-called "nonstandard", semantics for $L2K$ and $L2K-$, the models for $L1K$ and $L1K-$ are qualified in the range of values of the predicate and function variables. In particular, a Henkin model is a 4-tuple $\langle d, D, F, I\rangle$ in which, for each n, $D(n)$ is a nonempty set of subsets of d^n, $F(n)$ is a nonempty set of functions from d^n to d, and d and I are as before (except that what I assigns to n-place predicate and function constants in K must be restricted to $D(n)$ and $F(n)$, respectively). It is now just $D(n)$, and not the full Cantorian powerset of d^n, that is the range of the values of the n-place predicate variables; and, similarly, it is $F(n)$, and not d^{d^n}, that is the range of values of the n-place function variables. No constraints are imposed on $D(n)$ and $F(n)$ other than that they be nonempty, but, as I have noted elsewhere, even this condition can be dropped in the logic of natural realism, where n-place predicate variables are taken to range over natural properties (for $n = 1$) and relations (for $n > 1$) – and where function variables are simply deleted. Whether or not there are any natural properties or relations is a strictly empirical, and not a logical matter; and, independently of such empirical considerations, there could be natural properties, but no natural relations, or there could be natural binary relations but no natural properties or n-ary natural relations, for $n > 2$, and so on for other combinations as well. The logic of natural realism is in this regard a "free" logic – i.e. a logic "free of existential presuppositions" – with respect to the values of the bound predicate variables. Of course, this is not the sort of semantics we are interested in for strictly mathematical theories, nor for the analysis of intensional discourse or the semantics of natural language in general.

In what Henkin called "general" models, the sets $D(n)$ and $F(n)$ are constrained under the condition that all instances of the comprehension principle for concepts and functions must be satisfied. If they are further constrained to satisfy the axiom of choice as well, then the resulting models are what Shapiro calls *faithful to* $D2$ (or $D2-$). A special case of faithful models are the *full* Henkin models in which $D(n)$ is the full Cantorian power-set of d^n and $F(n)$

is d^{d^n}, which are equivalent to the so-called "standard" models $\langle d, I\rangle$. The semantics of "standard" models, in other words, amounts to a special case of the semantics of Henkin models, and the "standard" semantics of $L2K$ amounts to a restriction on the more general Henkin semantics of $L2K$. The price of the restriction, as Shapiro observes (though only in the terminology of "standard" semantics), is (1) that no deductive system can be complete with respect to it (i.e. under the "standard" semantics), (2) that the (upward and downward) Löwenheim-Skolem theorems fail for it, and (3) that it is not compact for $L2K$ (86f).

With respect to "general" and faithful models, on the other hand, Henkin showed that the deductive systems $D2$ and $D2^*$ are sound, complete, compact, and satisfy the Löwenheim-Skolem theorems – just as $D1{=}$ is sound, complete, compact, and satisfies the Löwenheim-Skolem theorems with respect to the first-order semantics for $L1K{=}$. The problem, Shapiro argues, is that in the case of non-full Henkin models, this semantics involves a restriction on the notion of *all subsets* of d, the domain of discourse, as values of monadic predicate variables, and, similarly, a restriction on the notion of *all subsets of* d^n (or *all* n*-ary relations* in the set-theoretic sense) as values of the n-place predicate variables (and a similar restriction in the case of the values of function variables). For Shapiro, "once a domain (for the first-order variables) is fixed, there is a reasonably clear and unambiguous understanding of such locutions as 'all relations [in the set-theoretic sense]' or 'all subsets' thereof" (204) – as determined in the metalanguage by the Cantorian notion of powerset, even in the case of an infinite domain.

The reason for this insistent reminder that the Cantorian notion of powerset is involved in Shapiro's notion of *all subsets* of an infinite domain is that, as Shapiro himself observes, there are logicians for whom "there is no clear understanding of the totality of the subsets of d (i.e. the powerset of d)" (247) in the case of an infinite domain, as understood under Cantor's powerset theorem. "The independence results in set theory suggest this. The powerful axioms of ZFC do not suffice to fix the powerset of the set of natural numbers, the simplest infinite powerset" (ibid.). There is, moreover, an alternative framework (or really several such) based on what I have called Frege's double-correlation thesis in which Cantor's powerset theorem, at least as applied to the total universe of discourse, is refuted, and yet in which second-order statements of the form, $\forall X \exists Y \forall Z \Phi$, which Shapiro takes as expressing our grasp of the locution 'all subsets' (247), are meaningful.[1] In this framework, it is concepts, and not "sets", or "subsets" of a domain, that are taken as the values of predicate variables, and classes (to the extent they are admitted at all) are assumed, as Frege put it, to have their being in the concepts whose extensions they are, and not (as on the Cantorian, iterative notion of set) in their members. A set-theoretic semantics for such a framework – which, because concepts are not sets, cannot but be semantics from an "external" point of view – cannot be based on the

[1] See "Cantor's Power-Set Theorem Versus Frege's Double-Correlation Thesis", *History and Philosophy of Logic*, 13 (1992), 179-201.

Cantorian notion of the subsets of d (or of the subsets of d^n), the domain of discourse, as values of the (n-place) predicate variables, except at best in the sense of a conceptual manque in place of the role of concepts. It is just in the sense of such a conceptual manque that the sets in $D(n)$ and $F(n)$ of a Henkin model are to be understood.

A set-theoretic (and therefore strictly "external") semantics corresponding to this approach is described by Shapiro (who notes (24) that it is close to a semantics described elsewhere by this reviewer[2]). On this semantics for $L2K$ (and $L2K-$), which Shapiro calls "first-order semantics" (a terminology that this reviewer finds misleading and question-begging[3]), a model is a 4-tuple $\langle d, d_1, d_2, \langle I, p, a\rangle\rangle$, in which d (the domain of discourse, or range of values of first-order variables) is a nonempty set, I is an "interpretation function assigning items constructed from d" (74) to the symbols in K; and, for each n, $d_1(n)$ and $d_2(n)$ are nonempty ranges of values for the n-place predicate and function variables, respectively; and, also for each n, $p(n)$ is a subset of $d^n \times d_1(n)$ and $a(n)$ is a function from $d^n \times d_2(n)$ to d. (The idea is that $p(n)$ is the interpretation of the n-place "predication" relation and $a(n)$ is the interpretation of the n-place "application" function.) A Henkin model $\langle d, D, F, I\rangle$, as Shapiro notes, "simply *is* the first-order model $\langle d, d_1, d_2, \langle I, p, q\rangle\rangle$ such that d_1 is D, and F is the 'real' application function from the various $F(n)$ and d^n to d" (75). "In short, Henkin semantics and first-order semantics are pretty much the same" (76), regardless whether or not additional constraints, such as satisfaction of the comprehension principle, are imposed on them.[4]

As construed by Shapiro, Henkin and "first-order" semantics are "nonstandard" because "the range of the predicate variables varies from model to model, even if the range of the first-order variables is held fixed" (173); and, therefore, according to Shapiro, whereas in "standard" semantics, "predication, or membership, is logical" (6), in Henkin and "first-order" semantics, predication or membership is nonlogical – as determined of course by the set-theoretic metalanguage based on the Cantorian notion of powerset. In other words, as the membership relation characterized by the Cantorian notion of set and powerset, predication, according to Shapiro, is a logical relation, whereas any purported predication relation that is not in accord with the Cantorian notion is nonlogical. The idea is that expressions for logical concepts (including predication or membership) have a fixed interpretation (with respect to the set-theoretic back-

[2] "Predication versus membership in the distinction between logic as language and logic as calculus", *Synthesé*, 77 (1988), 37-72, especially section 6. The models of this paper go beyond Shapiro's version by including a set of possible worlds and a Fregean correlation of concepts with objects in the domain, which is used to interpret nominalized predicates as abstract singular terms.

[3] "Standard" second-order semantics can be formulated in first-order set theory no less so than Henkin or this so-called "first-order" semantics, and therefore, from that perspective, each can be viewed as representing a muti-sorted first-order logic no les so, or more so, than the other.

[4] It should be noted, however, that these so-called "first-order" models can be extended to modal and intentional contexts (with or without nominalized predicates as abstract singular terms) in which even necessarily co-extensive concepts might not be identical. Such an extension does not seem possible for Henkin models.

ground), and that, relative to this fixed interpretation, the invariance of logical truth (or validity) is determined only by what is invariant across all domains of discourse. It is for this reason Shapiro can claim that "in fixing a domain one thereby fixes the range of both the first-order variables and the second-order variables" (73) – as determined by the Cantorian notion of powerset – and that "[t]here is no further interpreting to be done" (ibid.).

On the alternative, so-called "nonstandard", view, which is really based on the priority of concepts over classes, the invariance of logical truth (or validity) is an invariance not only with respect to all domains, but also with respect to all of the ways of conceptually structuring those domains as determined by the laws of compositionality for concept-formation (which is also where the role of the fixed interpretation of logical constants comes in). It is concepts, and not sets, that are predicable entities (and that underlie the correct use of predicate expressions and predication in language), and the primary constraint that has to be imposed on any "external" set-theoretic modeling of such predicable entities is that the laws of compositionality for concept-formation (which is represented here by the comprehension principle, (CP)) must be satisfied in all such models (i.e. that the models be "general" in Henkin's sense). It is this notion of logical truth (or validity) that coincides with Frege's "laws of logic" in his *Begriffsschrift*, and that coincides with provability (or deductive consequence) in $D2^*$, and which therefore is compact and satisfies the Löwenheim-Skolem theorems.

The main drawback on these results, Shapiro notes, is that "no theory with an infinite Henkin model is what may be called 'Henkin-categorical' " (95), i.e. "second-order languages with Henkin (or first-order) semantics are not adequate to characterize infinite structures up to isomorphism" (ibid.) – assuming, to begin with, that infinite structures can be characterized up to isomorphism (and are not subject to Skolem relativism).

2 Second-Order Versus First-Order Logic

Categoricity is an important component in Shapiro's argument for second-order logic under its "standard" semantics and against first-order logic. For example, where $A = \{0, s, +, \cdot\}$, first-order arithmetic based on $L1A$= (and consisting of successor, addition, and multiplication axioms, and an induction schema) is not categorical, i.e. not all models of first-order arithmetic are isomorphic. By comparison, all models of second-order arithmetic, AR, consisting of the conjunction of the successor, addition, and multiplication axioms, and the induction axiom,

$$\forall X[(X0 \wedge \forall x(Xx \to Xsx)) \to \forall x Xx],$$

which is a formula of $L2A$, are isomorphic with respect to the "standard" semantics of $L2A$. Similarly, first-order real analysis based on $L1B$, where $B = \{0, 1, +, \cdot, \leq\}$, is not categorical, i.e. not all of models of first-order real analysis are isomorphic; and yet, based on $L2B$, all models of second-order real analysis, AN, which has a second-order axiom of completeness, are isomorphic under the "standard" semantics of $L2B$.

The language of set theory has the membership symbol, $\in$, as its only non-logical constant. Where $Z2$ is the conjunction of the first-order axioms of ZF, Zermelo-Fraenkel set theory, and the second-order axiom of replacement,

$$\forall f \forall x \exists y \forall z (z \in y \equiv \exists w (w \in x \wedge z = fw)),$$

then the result of adding $Z2$ to $D2$ is a deductive system equivalent to Morse-Kelly set theory. As Shapiro observes, if M is a "standard" model of $Z2$, then M is isomorphic to an inaccessible rank, which means that if $M1$ and $M2$ are both "standard" models of $Z2$, and $M1$ is not isomorphic to $M2$, then "one of them is isomorphic to an 'initial segment' of the other" (86), a result indicating that, unlike first-order set theory, second-order set theory is "almost categorical".

The second-order theories of arithmetic and set theory are not conservative extensions of their first-order counterparts. In particular, a "truth definition" for the first-order theories can be formulated in the second-order theories, and thus one can prove the consistency of the first-order theories with respect to their second-order counterparts. More importantly, there are "nonstandard" models of the following sets of first-order formulas:

$$AR - 1: \quad \{\Phi : \Phi \text{ is a first-order formula of } AR \text{ and } \models AR \rightarrow \Phi\},$$

$$AN - 1: \quad \{\Phi : \Phi \text{ is a first-order formula of } AN \text{ and } \models AN \rightarrow \Phi\},$$

$$Z2 - 1: \quad \{\Phi : \Phi \text{ is a first-order formula of } Z2 \text{ and } \models Z2 \rightarrow \Phi\},$$

which means that there are uncountable models of $AR - 1$, countable models of $AN - 1$, and countable models of $Z2 - 1$. This is the situation known as Skolem's paradox, which is a consequence of the fact that, unlike their second-order counterparts, the first-order theories of arithmetic, real analysis, and set theory are not categorical.

There are notions other than categoricity that are also indicative of the difference between the expressive powers of second-order and first-order logic. There is a second-order formula, $INF(X)$, for example, that asserts there is a one-to-one correspondence from X to a proper part of the extension of X, i.e. that X is (Dedekind) infinite, in which case $FIN(X)$, defined as $\neg INF(X)$, says that X is finite. Both of these are formulas of the pure second-order language $L2$, and they are satisfied only in models having an infinite, or finite, domain, respectively. The negations of first-order formulas that are satisfied only in an infinite domain do not similarly guarantee that a domain in which they are satisfied must be finite. The notion of finitude, in other words, "cannot be captured in a first-order language, any first-order language" (101), and yet, as Shapiro notes, "[f]initude is a widely used notion, occurring throughout mathematics" (102), which means that "the language used to formalize mathematical practice should be capable of expressing this notion" (ibid.), and hence at least second-order.

Other cardinality notions, such as the cardinality of X *being less than or equal to* the cardinality of Y, and the cardinality of X *being equal to* the cardinality of Y, are also expressible in the pure second-order language $L2$, which means that the Schröder-Bernstein theorem can be stated in $L2$ – and, as Shapiro notes, proved in $D2$ (and $D2^*$ as well) (102f). The notion of *being countable* is expressible in $L2$ in terms of the notions of finitude and equi-cardinality, and *being denumerably infinite* is then expressible in terms of being countable and infinite, which then can be used to express being ALEPH-0, and so on to ALEPH-1, ALEPH-2, etc. The continuum hypothesis, CH, can be formulated in $L2$, the language of pure second-order logic, and, as Shapiro notes, the following sentence,

$$NCH: \quad \forall X(ALEPH-1(X) \rightarrow \neg CONTINUUM(X)),$$

is logically true (under the "standard" semantics) if, and only if, the continuum hypothesis is false, and therefore either CH or NCH is logically true (under the "standard" semantics). On the other hand, standard forcing techniques have been applied to second-order set theory using Henkin models, showing that neither CH nor NCH (nor their negations) can be deduced in $D2$ (or $D2^*$) from $Z2$ (second-order ZF), and therefore neither can be deduced from $D2$ (or $D2^*$) alone. This situation, Shapiro observes, "is an instance of the incompleteness of second-order logic. One of CH, NCH is an unprovable logical truth" (105) – that is, logically true in the sense of the "standard" semantics".

The notion of the *well-foundedness* of a relational concept E can be expressed in $L2$ in terms of a pure second-order formula,

$$\forall X[\exists x Xx \rightarrow \exists x(Xx \wedge \forall y(Xy \rightarrow \neg Eyx))],$$

but, as noted by Shapiro, it "cannot be characterized in a first-order framework" (108). Similarly, the notion of a *well-ordering R of the instances of a concept X*, $WO(R, X)$, can also be expressed in $L2$, and yet "[t]here is no adequate formulation of this notion in a first-order language except in the trivial cases in which the extension of X has a fixed finite bound on its cardinality" (106) – the reason for this being the compactness of first-order logic. The compactness of $D2$ and $D2^*$ does not affect the characterization in $L2$ of well-foundedness, however; nor does it affect the expressibility of the well-ordering principle,

$$WOP: \quad \forall X \exists R(WO(R, X)),$$

as a pure second-order formula of $L2$. But then, despite the presence of a global axiom of choice, AC, in $D2$, a faithful Henkin model for $D2$ shows that WOP is not a theorem of $D2$. The converse does hold, however; in particular, it can be shown that $WOP \rightarrow AC$ is provable in $D2^*$ (107f).

A common mathematical practice noted by Shapiro is the definition of a set in terms of closure conditions on an initial or *basic* set B and a collection H of operations or relations; e.g., M is the *minimal closure of B under H* if M is the smallest set containing B and closed under the operations and relations in H. Dedekind's definition of the natural numbers is by means of a minimal closure, and so is Frege's definition in terms of the ancestral of the predecessor relation. Indeed, the proofs of categoricity of AR and AN are based on several such minimal closures. As Shapiro notes, the notion of an object x being in the minimal closure of a concept Y under a relation R can be expressed in $L2$ as follows:

$$MC(x, Y, R): \ \forall X[(\forall y(Yy \to Xy) \wedge \forall y \forall z((Xy \wedge Ryz) \to Xz)) \to Xx].$$

That such a minimal closure exists can be proved in $D2^*$ by means of an (impredicative) instance of the comprehension principle, (CP),

$$\exists Z \forall x(Zx \equiv MC(x, Y, R)).$$

For first-order logic, on the other hand, it can be shown by means of a compactness argument that "no collection of first-order formulas can successfully characterize any non-trivial minimal closure" (99), which is further evidence, Shapiro notes, that "first-order languages are inadequate" (ibid.).

Note, however, that the minimal closure that can be proved to exist with respect to a concept Y and relation R is only a concept and not a set as the correlate or extension of the concept. This is important because in set theory, which is the basis of Shapiro's "standard" models, such minimal closures generally turn out to be "too big" to exist as sets, as opposed to proper (or ultimate) classes (which, in effect, amount to concepts that have no sets corresponding to them). The construction of an appropriate set then has to be done "locally" inside one of the values of 'X' in the definition of minimal closure. This is not the case in the alternative frameworks mentioned in section one, where Cantor's powerset theorem is disproved and the Boolean structure of concepts holds not only for concepts (as values of predicate variables) but also for their extensions (as values of first-order variables) as well.

In such a alternative framework, nominalized predicate expressions are allowed to occur as abstract singular terms – the way they are in type theory, except in this case predicates are not typed (other than in the sense of their degree or ad-icity), and, assuming extensionality (Frege's basic law Va), they denote not the concepts that predicates stand in their functional role as predicates, but the extensions (or intensions, if extensionality is dropped) of those concepts instead. Cantor's powerset theorem is expressed in such a framework in either of several equivalent forms, but any one of which must have a nominalized predicate occurring as an abstract singular term – such as the simplest version, which is formulated as follows:

$$CT: \quad \forall R[R \ is \ many - one \ \to \exists X(\forall x[Xx \to \exists y R(x, y)] \wedge \neg \exists x R(x, X))].$$

This is noteworthy because Shapiro claims that Cantor's theorem can be expressed in $L2$, where predicates occur only as predicates and not also as abstract singular terms. Shapiro's formulation is

$$\forall R \exists X \forall x \exists y[(R(x, y) \wedge \neg Xy) \vee (\neg R(x, y) \wedge Xy)],$$

which does have a superficial similarity to Cantor's theorem, but which also is trivially equivalent to $\forall R \exists X \forall x \exists y[Xy \equiv \neg R(x, y)]$. This last formula, however, is derivable from

$$\forall R \exists X \forall x[Xx \equiv \neg R(x, x)],$$

an instance of the comprehension principle, (CP), and therefore it is provable in $D2^*$. But $D2^*$ (extended to include nominalized predicates as abstract singular terms) is the underlying logic of these alternatives (known as λHST^* and HST^*_λ) where the above and other equivalent versions of Cantor's theorem are refuted (directly in λHST^* and in slightly modified forms in HST^*_λ). In other words, given the refutation of the most explicit versions of Cantor's theorem in a framework having $D2^*$ as its underlying logic (and equiconsistent with weak Zermelo set theory or the theory of simple types), the comprehension principle of $D2^*$ (with or without nominalized predicates as abstract singular terms) cannot really be said to contain Cantor's theorem as an instance, contrary to what Shapiro claims on this matter.

3 The Rejection of Epistemological Foundationalism

By foundationalism, Shapiro means "the view that it is possible and desirable to reconstruct each (legitimate) branch of mathematics on a completely secure basis, one that is maximally immune to rational doubt" (25). This is an epistemological doctrine, and as such, Shapiro observes, it "has few proponents today, and for good reason" (vi), which is why he urges "a thorough rejection of foundationalism" (ibid.). In particular, it is his view that "just as we have learned to live with uncertainty in virtually every special subject, we can live with uncertainty in logic and foundations of mathematics, and we can live well" (25).

Two foundationalist efforts discussed by Shapiro are logicism, especially as held by Frege and Russell, and Hilbert's formalist program, for which "only a certain 'finitary' core of mathematics is meaningful, even though most branches go well beyond this" (28). Frege did maintain a certain form of epistemological foundationalism, but, as Shapiro himself notes, he also "had an ontological orientation as well" (54) in which numbers in particular are "logical objects" – namely, classes that have their being in the concepts whose extensions they are, as opposed to sets in the Cantorian sense. In general, an ontological foundationalism need not be committed to an epistemological foundationalism, and, as in conceptual realism, one might adopt something like a Fregean approach without assuming epistemological foundationalism.[5] Russell, incidentally, did not argue for certainty in logic and mathematics, but maintained instead that "[o]ur reasons for believing logic and pure mathematics are, in part, only inductive and probable in spite of the fact that, in their *logical* order, the propositions of logic and pure mathematics follow from the premises of logic by pure deduction."[6]

The main point about logicism is that the pure formal language of logic

[5] A brief description of conceptual realism is given in the article cited in footnote 1. A more detailed version is described in "Conceptual realism versus Quine on classes and higher-order logic," *Synthese* 90 (1992), 379-436, especially section 8.

[6] "Logical Atomism" (1924), reprinted in *Logic and Knowledge*, edited by R.C. Marsh, London: George Allen and Unwin LTD, 1956, p.326.

– which always is at least second-order and has nominalized predicates as abstract singular terms either on a par with other singular terms as values of the first-order variables (as in Frege's case, and as originally conceived by Russell as well), or, as in type theory, as values only of predicate variables of the same logical type – is a completely interpreted language that can be applied to any subject matter whatsoever. Such a logical language is to contain not only a *calculus ratiocinator*, i.e. a deductive system, but a *lingua characteristica* as well, i.e. a language for all rational, scientific discourse. "There is no independent metatheory" (28), Shapiro notes, of such a view of logic as language – but that does not mean that an "external" set-theoretic semantics (and in particular a Henkin or "first-order" semantics) cannot be used as a guide in the determination of logical truth or validity.

Neither logicism nor formalism were successful as foundationalist programs. In logicism, in particular, classes in the logical sense (i.e. as extensions of concepts), gave way to classes in the mathematical sense, i.e. sets in the sense of the iterative hierarchy as based on Cantor's theorem. But, aside from the independent development of set theory by mathematicians, that was mainly because of the contradiction in Frege's theory as a result of Russell's paradox on the one hand, and because of the unnaturalness of type theory as a model of natural language on the other. It was not, in other words, because of any epistemological doctrine associated with logicism. (Frege's logic, incidentally, can be reconstructed in terms of either of the systems λHST^* and HST^*_λ, which do provide natural models of natural language, but which also can serve as a foundation for conceptual realism.) In any case, logicism certainly cannot be accused of psychologism, "the view that the aim of logic is to characterize a certain kind of subjective certainty" (31). Indeed, Shapiro notes that "[m]ost foundationalist authors reject psychologism," and claim instead that "logic deals with objective norms of actual correct inference" (ibid.). This need not mean that every deducible argument in the logicist's system corresponds to a correct inference in natural language; for, as a complex inferential engine the system may validate arguments that do not arise in natural language except under the artificial conditions of a formal system. But the question whether the system is "*exhaustive* in the sense that every correct inference of the natural language corresponds to a deducible argument in the formal language" (33) is another matter.

According to Shapiro, if logicism were either (1) "to make a concession to psychologism and assert that the inferences that correspond to the deducible arguments all *seem* certain, or are obvious", or (2) "postulate a faculty of normative epistemic judgement, a faculty outside the purview of psychology", then, it cannot "assure us that the system is exhaustive" (ibid.). On the other hand, if "the logicist can live with the possibility that the deductive system may have to be expanded one day", that would be "an incompleteness of sorts" (34). Such an "incompleteness" could be with respect to an expanded categorial analysis of natural language, however, i.e. in the recognition over time of new logical categories corresponding to certain types of expressions of natural language, so that certain correct inferences of natural language that could not be explained before come to be covered under the expanded theory of logical forms. Whether ultimately such a process of extending the theory of logical forms must result in an "exhaustive" system is a question that can be left open, subject to further assumptions and investigations of natural language. In any case, the logicist need not agree that the "incompleteness" in question is a result of what Shapiro calls "standard" semantics, i.e. that the system is committed to completely representing the "standard" logical truths of set-theoretic semantics as based on the Cantorian notion of powerset.

For Shapiro there are today two conceptions or "orientations toward modern logic" (35), namely, the *foundational* and the *semantic* conceptions. On the foundational conception one gives axioms and rules of inference for deducing theorems, and logic is "seen as the ideal of what may be called *relative justification*, the process of coming to know some proposition *on the basis of* others" (36). This sort of enterprise, according to Shapiro, is "a limited version of logicism" (ibid.), where the implicit assumption is that "our pre-theoretic ability to detect correct justification suffices to sanction counterparts of each immediate inference" (37). On the semantic conception, on the other hand, "validity is characterized in terms of models or interpretations of the language in question" (ibid.). Admittedly, this characterization depends on the distinction between logical and nonlogical constants (and hence on an implicit theory of logical form, such as is proposed in logicism?). Following Tarski, Shapiro thinks that this distinction need not "be determined in advance of logical theory", and "may hold for a natural language only relative to a model-theoretic semantics for a formal language that corresponds to it" (39). In other words, on this conception, "the plausibility of a proposed semantics depends on the extent to which the class of models corresponds to the intuitive notion of 'interpretation' or 'possible world', the one active in pre-formal judgements of validity" (ibid.).

One could adopt both a foundational and a semantic conception (and achieve a "marriage" of sorts) – as opposed to adopting one and rejecting the other – in frameworks where there is a convergence on a single consequence relation (as in second-order logic under Henkin or "first-order" semantics). Shapiro, however, opts for accepting the semantic conception and rejecting the foundational view. His main argument for this is that "the extension of the notion of 'ideal justification' may be open-ended" (44), i.e. "one should admit the possibility that we may one day discover arguments that conform to foundational standards of ideal justification, but have counterparts in [the given formal language] L that are not deducible in the deductive system. ... This is consonant with the essential incompleteness of the logic" (ibid.). But this "essential incompleteness" could be the result only of an "incomplete categorial analysis" of the logical forms underlying natural language, where the counterparts in L that are not deducible are the best we could do as analyses of those arguments, and not because the so-called "standard" models of set theory provide the "true" and "correct" interpretation of natural language. The convergence on a single consequence relation would then be only for that part of natural language covered by the logical forms adequately represented in the formal language L.

Shapiro rejects the extreme opposite of foundationalism, which "would make justification *completely* subordinate to semantics" (48). Instead, he adopts "a middle course in which we keep an 'intuitive' notion of justification and use it to guide our theorizing, both semantic and deductive, and to evaluate deductive systems" (ibid.), which seems a very reasonable approach. It should be noted, however, that although "[t]here is no problem here with justification being laden with theory, even being laden with set theory" (ibid.), it is not clear that the membership relation of set theory, as based on the Cantorian notion of powerset, can be the criterion by which to judge an adequate semantics for predication in natural language.

The way to understand natural language, including the informal discourse of mathematics, Shapiro agrees, is through the *Use Thesis*, which is a thesis not about meaning or semantics but about *understanding*. The idea is that one "understands the concepts embodied in a language to the extent that one knows how to use the language correctly" (211). That is, "the ability to use a language *constitutes* understanding and thus grasping the concepts" (ibid.). Indeed, in conceptualism, concepts are none other than the cognitive capacities, or cognitive structures otherwise based upon such capacities, that constitute our ability to use the different types of expressions of natural language. Thus, predicable concepts, for example, are cognitive structures based upon intersubjectively realizable capacities to identify, characterize, and relate objects in various ways; it is these concepts that underlie our ability to follow the rules of language in the use of predicate expressions. Indeed, strictly speaking, our knowledge of the rules of language is not propositional knowledge (knowledge-that), but knowledge in the sense of having concepts as cognitive capacities; and our following those rules in particular speech acts is just a matter of our exercising those concepts as cognitive capacities. It is concepts in this sense that are values of the predicate variables in conceptualism, and not sets.

Though Shapiro refers to concepts in his description of the Use Thesis, he takes only sets and never concepts as the values of predicate variables. Predicable concepts do have extensions (or at least most do), which can be taken as their surrogates in model-theoretic semantics. That is, as the determinant for the correct use of a predicate expression in different contexts (and thereby of the truth conditions for that expression), a predicable concept will in general determine a class as its extension (in such a context), and that class may then be taken as a proxy for the concept (as, e.g. in Henkin semantics). But it is an enormous jump from this notion of a class as the extension of a concept to the notion of a set as determined by Cantor's powerset theorem. Indeed, in conceptual realism (as represented in λHST^* or HST^*_λ), where nominalized predicates are allowed to occur as abstract singular terms, and where (assuming extensionality) such nominalized predicates denote the extensions of the concepts that predicates stand for in their role as predicates, Cantor's theorem is refuted (as already noted in section one).

The Use Thesis, Shapiro claims, "does not demand an anti-realist ontology nor a non-model-theoretic semantics" (212), contrary to what Dummett has claimed on its basis instead. Certainly, it does not demand anti-realism in conceptual realism – at least not in regard to the classes (i.e. extensions, or intensions if the extensionality axiom is rejected) that nominalized predicates are assumed to denote (as values of the first-order variables) — which is a paradigmatic framework for the Use Thesis. (Concepts do not exist independently of the human capacity for concept-formation, however, and perhaps in this sense the framework can be said to be "anti-realist" in regard to the values of predicate variables.) And model-theoretic semantics is acceptable so long as the models are appropriately constrained in accordance with the laws of compositionality for concept-formation (the way they are in "general" Henkin models). Given the Use Thesis, Shapiro observes, all that follows "is that the ontology and truth conditions of semantics must be compatible with the learnability and understanding of a language through its use. Truth conditions should not be that far removed from the knowledge implicit in the correct use of a language, and from whatever it is that one learns when acquiring facility with a language. There should be a natural (if not inevitable) link between use and truth conditions" (ibid.). But, again, it is a very big jump from the concepts (and their extensions) presupposed by the Use Thesis to the set-theoretic models of "standard" semantics as based on the Cantorian notion of powerset, and which is the basis of the categoricity results favored by Shapiro.

Indeed, Shapiro goes so far as to "propose that, on the Use Thesis, Skolemite relativism is to be rejected out of hand" (213), and that "the *use* of mathematical discourse ... is not captured by first-order languages or effective deductive systems, namely the sense in which the discourse outstrips, or even 'transcends', its previous manifestations" (ibid.). Certainly, in conceptual realism, "first-order languages are not adequate to capture important aspects of the *use* of the relevant languages" (ibid.), at least not with respect to the concepts presupposed by that use, i.e. the concepts that are the basis of predication in any use of language (and for whatever purpose). But being at most potentially infinite (i.e. countable), such concepts cannot be the basis of any categoricity results for infinite domains; and if classes are assumed to have their being in the concepts whose extensions they are, then, rather than reject Skolemite relativism, the Use Thesis is committed to it – but in a form appropriate to second-order logic under its Henkin or "first-order" semantics. Semantics, as Shapiro himself admits, "comes after the original natural language of mathematics is understood, and the original understanding does not consist of grasping a model, intended or otherwise" (213), such as that of the natural or real numbers. "Understanding the language is knowing how to use it" (ibid.), which suggests that the Use Thesis, especially as developed in conceptual realism, provides a conceptual or philosophical rationale for Skolemite relativism, rather than a reason for rejecting it.

4 Will the Real Logical Classes (Sets?) Please Stand Up

An important component of Shapiro's "standard" semantics is what he calls the logical notion of set, by which he means the notion of a subset of a universe of discourse in the Cantorian sense, and not the logical notion of a class as the extension of a concept. To be sure, logical sets are like classes in the logical sense in that, unlike iterative sets, both "exemplify what may be called a Boolean structure" (177), i.e. both have complements and there is a universal "set" or class. The main difference is that whereas Cantor's theorem is assumed to hold for logical sets, it fails to hold for classes in the logical sense (as represented in λHST^* or HST^*_λ).

Another difference has to do with the fact that "[t]here are no logical sets *simpliciter*, only logical sets within a given context" (18) – because, according to Shapiro, "the word 'set' is like an indexical expression of ordinary language" (ibid.). This way of phrasing the issue cannot be right as it stands, however, because otherwise there would be no him, her, them, I, thou, or us, *simpliciter*; that is, it cannot be right because, as just expressed, we are objects referred to by indexical expressions, and yet we exist regardless of whether or not those expressions are used in a different context to refer to other people. Objects do not just pop into being when referred to by an indexical expression, and then pop out of being when the context is changed and the expression is used to refer to a different object. Indexicality simply does not have that kind of ontic effect or significance.

The point about indexicality, according to Shapiro, is that a logical set is a subclass of a universe of discourse, which means that it "always involves a domain fixed by context" (177). Putting it this way comes very close to what is meant in conceptualism by the claim that reference to the subsets of a domain of discourse of a set-theoretic model, as opposed to reference to the concepts that are internal to a theory regarding that domain, is at best reference with respect to an external semantics, and that the iterative sets serving as proxies for concepts in that external semantics cannot really be those concepts. Similarly, just as the domain of a "standard" model cannot itself be a "logical set", no less *the* universal logical set (there being none *simpliciter*), but only an iterative set going proxy for such, so, too, the subsets of that domain cannot be "logical sets" (again, because there are none *simpliciter*), but only iterative sets serving as proxies for such. Or, just as the domain of the model is not universal with respect to the metatheory, so, too, none of the subsets of that domain have complements with respect to the set theory of the metalanguage. Thus, just as one might say there are no concepts *simpliciter*, other than the concepts that can be realized within the context of a theory (of ordinary or scientific language), so too "there are no logical sets simpliciter, only logical sets within a given theory" (177).

So much for similarity, however, because a crucial difference remains. In particular, logical sets, even if only taken as values of predicate variables within a theoretical context, are objects, whereas concepts, as rule-following cognitive capacities, are not objects but *unsaturated* cognitive structures in a sense analogous to, but not the same as, what Frege meant by an unsaturated function.[7] That is, as rule-following cognitive capacities that can be realized in a theoretical context (of ordinary or scientific language), concepts can be exercised by different people at the same time, as well as by the same person at different times – or, they may, relative to that theoretical context, never in fact be exercised at all. It is this nonoccuurent, or purely dispositional, status of concepts as rule-following cognitive capacities that need not be exercised at any given time, or even ever, in a given theoretical context that explains in part their unsaturated nature, i.e. why concepts cannot be objects. Another part of that explanation is that the exercise, or saturation, of such a capacity in a speech act of the theoretical context in question is what informs that act with a predicable nature, and it is this aspect of a concept that is the basis of predication in that context.

Nothing at all like this holds for what Shapiro calls logical sets, and one can only wonder what it is about them, or the "logical" relation of membership with respect to them, that explains predication in language and thought. That they are objects, and not, e.g., unsaturated functions from objects to truth values, as in Frege's theory, is implicit throughout this book. But, even if one were now to assume otherwise, i.e. reconstrue logical sets as Fregean *Begriffe*, there would still be the problem of Cantor's theorem; for, whereas Cantor's powerset theorem is assumed to hold for logical sets, the opposite is assumed by Frege in terms of his double-correlation thesis. A similar observation applies if one were to assume that logical sets are the extensions of Fregean *Begriffe* – or the extensions of concepts as cognitive capacities. In any case, Shapiro would reject such a reconstrual as part of the "foundational" conception of logic. "The present study," he makes clear, "eschews foundational questions, such as whether the classes in the range of our second-order variables, and the concomitant membership relation, are ultimately based on concepts, or whether the classes are based on the iterative notion of set" (20). And yet, as logical sets, there really are no such classes *simpliciter*.

Although he admits that first-order set theory "provides a uniform semantics for all (or almost all) formal languages" (251), and that 'the notion of 'second-order logical truth' is set-theoretically definable" (163) – indeed, that the "standard" semantics of second-order logic is "not quite as rich and complex as the notion of (first-order) set-theoretic truth" (ibid.) – Shapiro is not also willing to admit that logical sets are none other than the iterative sets of the background first-order set theory. This is because the semantics of first-order set theory cannot "fix the subject matter of the theory" (252) – i.e. the subject matter in the sense of the "standard" models of second-order set theory, which,

[7] Frege's *Begriffe* are really properties and relations of a form of logical realism, and not the cognitive structures of conceptualism. Unlike concepts as cognitive structures, there is no reason why there could not be uncountably many such independently real entities in such a framework as values of the predicate variables (despite the fact that Cantor's powerset theorem remains refuted by Frege's double-correlation thesis).

at least internally, according to Shapiro, assumes that the values of the predicate variables are logical sets. "There are models [of first-order set theory], for example, in which the membership relation is not well-founded, and by hypothesis, we have no principled way to 'rule out' the unintended ones" (ibid.). The result, in other words, is "another variant of the Skolem paradox, now applied to the background set theory" (252), which then leads to "an unavoidable relativity of virtually all mathematical notions" (253), which Shapiro maintains is not reflected by mathematical practice. But, as based on the Use Thesis, it is hard to see how mathematical practice can really settle this.

Another option is to regard the background first-order set theory as having "an *intended interpretation*" (253), just the way "the natural number structure" and "the real number structure" are taken as the intended interpretations of arithmetic and analysis. The fact that the background set theory has unintended models is then taken as irrelevant, because (shades of logicism) the theory is to be regarded as "fully interpreted" (ibid.). The problem with this option is that it leaves as a mystery how the intended model is "grasped, understood or communicated," which "is all the more perplexing since m [the intended model] is so complex. Without an independent characterization of m, it is not clear how the language of set theory overcomes the problem with characterizing structures in first-order languages" (254). It also leaves unexplained how predication in the background theory is to be explained.

A final option is to regard set theory as a "regimented version of the natural language of mathematics normally used as metalanguage, the 'mother tongue' of logical theory" (ibid.). Such a regimented version can have an "explicit syntax and a uniform ontology", which can be employed "without apology", because "[t]here is no perspective outside this language from which to discuss its interpretations, or its models" (ibid.). This option is "unobjectionable", according to Shapiro, because given "the overriding anti-foundationalism, it does not follow that second-order logic is undermined" (255). That may be true, but, by second-order logic, Shapiro means here only second-order logic as characterized in "standard" models, where it is the Cantorian notion of subset that interprets how "one can speak coherently of 'all subsets' of d [the domain of discourse] and 'all relations' on d" (ibid.) – an interpretation that is based on the logical notion of set, as opposed to the logical notion of class as the extension of a concept underlying the use of the natural language of mathematics.

5 Concluding Remarks

There is a great wealth of material about second-order logic in this book, only some of which we have touched on here. Chapter 6, for example, explains in a clear and precise way the "sense in which n-th-order logic, for $n \geq 3$, is reducible to second-order logic" (134), and there is an excellent section on reflection principles and the realm of so-called small cardinals in the set-theoretic hierarchy, as well as a section on "large large" cardinals. Three historical items are taken up and discussed expertly in chapter 7; namely, (1) the emergence of first-order logic (and semantics) as the “standard” in logic; (2) the development of set theory and the emergence of first-order ZFC; and (3) the contemporary controversy over the status of second-order logic. A fascinating debate that leads to a regress over the question of whether or not there is an unequivocal understanding of ‘all relations’ or ‘all subsets’ is described in chapter 8. Overall, it is really an excellent book, and, notwithstanding the philosophical and/or conceptual difference of opinion this reviewer has regarding the significance of “standard” as opposed to “nonstandard” semantics for second-order logic, the author is to be commended for a job well-done.