Notre Dame Journal of Formal Logic
Volume 40, Number 2, Spring 1999

Book Review

Penelope Maddy. *Naturalism in Mathematics*. Clarendon Press, Oxford, 1997. viii + 254 pages.

1 Theory choice in set theory Set theorists investigate sets. They also investigate theories. Set theorists investigate certain formal theories because:

1. their study of those theories teaches them interesting things about their own informal set theorizing; and
2. the metatheory of those theories is mathematically rich and interesting in its own right.

They get (1), of course, because theories like ZFC are successful formalizations of informal set theorizing. But they get (2) partly because certain features of informal set theorizing are omitted from the formalizations. Set theorists speaking a dialect of English can, for example, fully characterize the order type of the natural numbers. But it is the formalizations that *cannot* characterize ω categorically that are most attractive metamathematically. So there is a tradeoff: we get the most bountiful metatheory only if we concentrate on theories too weak to capture every aspect of informal set theorizing. First-order languages lack the full expressive capacity of informal mathematical discourse, but they have a metatheory that many mathematicians find stimulating and suggestive and entertaining and useful and generally wonderful. So three cheers for first-order languages! Let us just not forget their limitations.

The formal set theory of our dreams would have a great metatheory *and* would fully capture every important aspect of informal set theorizing. Well (consarnit!) there could be a perfect fit between the formal and the informal if the set theorists would limit their *informal* discourse. If only they would steer clear of those pesky informal locutions that are not firstorderizable! If only we could *make* them speak "an austere sublanguage of English that corresponds to the limited resources of first-order logic" (p. 212). Of course, we couldn't do so even if we tried. And if we did try, we ourselves would violate a fundamental principle of philosophical methodology and we would be urging the set theorists to erase a fundamental principle of set theoretic methodology.

The philosophical principle is MODESTY: " . . . philosophy is not in the business of criticizing and recommending reform of good mathematics on extra-mathematical grounds” (p. 171). More generally, “ . . . mathematics is not answerable to any extra-mathematical tribunal” (p. 184). A modest philosopher will, of course, not insist that informal mathematical discourse be emasculated.

The set theoretic principle is COMBINATORIALISM: “ . . . at each stage of the cumulative hierarchy [of sets], we include ‘all possible collections’ . . . of previously-formed entities.” That is, “ . . . we reject the requirement that all collections be definable [and] we also add a maximizing idea that every collection that can be formed at a given stage, will be formed at that stage” (p. 208). A combinatorialist believes that the iterative hierarchy is as thick as possible, that the power set operation delivers *all* the subsets of a given set. We might express this as follows. Things *form a set* just in case there is a set whose members are exactly *them*. Bach, Beethoven, and Brahms form a set just in case there is a set whose members are exactly Bach, Beethoven, and Brahms (that is, all three of them are members and every member is one of them). What, then, are the subsets of a set A? The combinatorialist says: *any* members of a set form a set; so *any* members of A form a set; and any set formed by members of A will, of course, be a subset of A. A set theorist confined to “an austere sublanguage of English that corresponds to the limited resources of first-order logic” could not say what I just said or anything equivalent to it and so could not even *express* a full-blooded combinatorialism. Our philosophically immodest linguistic revolution would prevent set theorists from expressing one of the key ideas behind a mathematically successful research program.

At the risk of sounding philosophically immodest, we might remark that the linguistic revolution would be a bad move even if it was the set theorists’ idea: “ . . . contemporary pure mathematics is pursued on the assumption that mathematicians should be free to investigate any and all objects, structures, and theories that capture their mathematical interest . . . . If mathematics is to be allowed to expand freely in this way, and if set theory is to play the hoped-for foundational role, then set theory should not impose any limitations of its own” (p. 210). Set theorists should, by all means, study the formal theories that have a nice metatheory. But we would hope that this leads to no artificial limitations on their informal set theorizing. Informal mathematical discourse “should be as powerful and fruitful as possible.” (Cf. p. 211.)

This idea that mathematicians should be unfettered has implications for theory choice. Suppose, for the moment, that our goal is to craft a theory that captures our conception of the set theoretic universe (as opposed to a theory that maximizes the fertility of our metatheorizing). Then some choices are clear-cut. Second-order Z (Z^2) is better than first-order Z because only the second-order version allows us to say how thick we think the universe is. (The ability of set theorists to make formal claims about the thickness of the universe should not be limited by their choice of formal vocabulary.) Z^2 is preferable to Z^2+ ‘there is exactly one limit ordinal’ because it is illiberal to place an artificial cap on the height of the universe. (In this context, categoricity is a vice.) Z^2 is preferable to Z^2+ ‘some Borel sets are not determined’ because the latter theory’s inconsistency follows from the consistency of second-order ZF (ZF^2). (Here and below, “consistency,” “consequence,” “implication,” etc. are to be understood semantically.) ZF^2 is preferable to Z^2 because the former incorporates an intuitive principle of set formation that yields an abundance of interesting and stimulating structures. (p. 211: “ . . . given that set theory is out to provide models for all mathematical objects and instantiations for all mathematical structures, one way in which it should MAXIMIZE is in the range of available isomorphism types.") ZF^2 is preferable to ZF^2+ 'there is no regular limit cardinal' because, once again, we should let the universe be as tall as it can be (and, as far as we know, the universe *can* be tall enough to accommodate a regular limit cardinal—indeed, ZF^2 plus a natural reflection principle implies that there are ω inaccessibles). ZF^2 is preferable to $ZF^2+ \sim$Con(ZFC) because the latter theory is known to be inconsistent.

Those choices are easy. Here is a harder one. Is there a measurable cardinal? ZF^2+ 'there is no measurable cardinal' appears to be limitative. (Let the universe be as tall as it can be!) But ZF^2 might actually *imply* that there is no measurable cardinal. Consider the claim that every set of natural numbers has a countable L-rank: $\wp(\omega) \subset \mathrm{L}(\omega_1)$. Call this '$\Psi$'. Either Ψ or $\sim \Psi$ follows from ZF^2. (ZFC proves: Ψ is true if and only if $\mathrm{H}(\omega_2)$ thinks Ψ is true. That is, to decide Ψ we need only look at sets hereditarily smaller than ω_2. Since all those sets live inside $\mathrm{V}(\omega_2)$, ZF^2 fully determines their structure. So ZF^2 implies either Ψ or $\sim \Psi$). If ZF^2 implies Ψ, then ZF^2 also implies that there is no measurable cardinal. In that case, the universe *cannot* be tall enough to accommodate a measurable cardinal and 'there is no measurable cardinal' is not really limitative at all.

On the other hand, any argument for the consistency of ZF^2+ 'there is a measurable cardinal' (ZF^2+ MC) is an argument that ZF^2 implies $\sim \Psi$. First-order derivations in ZFC + MC can supply at least two sorts of evidence for second-order consistency. The first is inductive: "No absurdities yet!" The second is conceptual: "We're getting a clearer idea of what this universe looks like and it seems coherent." Though such evidence is far from decisive, it is still worth having. So it would be imprudent to *ban* the exploration of ZFC + MC or to *require* that first-order set theorizing take place in ZFC + 'there is no measurable cardinal' since this would limit our access to important evidence about second-order consistency and consequence. Although the proposition 'there is no measurable cardinal' might place no real limit on the height of the universe, an exclusive commitment to this proposition might limit significant mathematical inquiry.

If $ZF^2+ V = L$ is consistent, then ZF^2 implies Ψ. First-order derivations in $ZF + V = L$ could supply evidence of the second, conceptual, sort for this consistency. (We do not need inductive evidence if we already believe in the consistency of ZFC.) First-order work inside L could help us to see whether $V = L$ is really compatible with our combinatorial conception of power set. So, as before, it would be imprudent to *ban* the exploration of $ZF + V = L$ or to *require* that first-order set theorizing take place in $ZFC + V \neq L$ or in ZFC + '$0^{\#}$ exists' since this would limit our access to important evidence about second-order consistency and consequence. Although $V = L$ might place a real limit on the height and thickness of the universe, an exclusive commitment to its negation might limit significant mathematical inquiry.

So ZFC + MC and $ZF + V = L$ are both worth investigating even though ZF^2+ MC and $ZF^2+ V = L$ cannot both be consistent. If we had to choose one or the other, though, $ZF + V = L$ seems the safer option. Even if $ZF^2+ V = L$ is inconsistent, the $\mathrm{L}(\alpha)$'s will still inhabit the universe of sets and theorems of $ZF + V = L$ will still characterize them. ZF^2 will imply Ψ^{L} even if it also implies $\sim\Psi$. More generally, every theorem of $ZF + V = L$ becomes a consequence of ZF^2 if we relativize its quantifiers to L. So none of our work will be lost even if $V = L$ bites the dust. ZFC $+$ MC is riskier. It may itself be inconsistent. And if $ZF^2 +$ MC is inconsistent, then the first-order theorems derived from MC tell us only what the universe would be like if it were inhabited by something impossible. Of course, we already know what that universe looks like! ZFC $+$ MC should certainly be explored. My point is only that this seems to involve a greater risk of wasting one's time.

I am assuming here an unwavering commitment to COMBINATORIALISM. If ZFC $+$ MC proved to be an extraordinarily attractive theory and if we somehow convinced ourselves of its consistency, then we might cease to care whether $ZF^2 +$ MC is consistent. This would require that we be willing to abandon the combinatorial conception of power set. (If $ZF^2 +$ MC turned out to be consistent, there would be no conflict with combinatorialism. But as long as the consistency question remains open, we would have to be open to the other alternative.) One would prefer, of course, that the combinatorial conception be replaced with one at least as well understood. People indifferent to the consistency of $ZF^2 +$ MC really ought to tell us what notion of power set they have waiting in reserve.

2 Maddy's model Though the above account seems a natural development of some central themes from *Naturalism in Mathematics*, it is not Professor Maddy's account. Maddy offers a model in which a theoretical entity known as a "naturalistic set theorist" attempts to choose rationally between competing set theories. Certain behaviors of the naturalistic set theorist (hereafter "NST") are identified as "the core of naturalized set theoretic activity" (p. 212). First: the NST asserts "versions of the axioms of Zermelo-Fraenkel . . . in an austere sublanguage of English that corresponds to the limited resources of first-order logic." Second: the NST uses informal logic to derive theorems from these axioms. Certain other behaviors are identified as "penumbral": for example, the NST is said to affirm COMBINATORIALISM and, allegedly, evaluates axiom candidates from the combinatorial perspective. More generally, when faced with an axiom candidate, the NST will survey penumbral principles in an effort to justify the candidate's inclusion in or exclusion from the core.

As an example of how the NST behaves, imagine its first encounter with Separation. If the NST were a genuine combinatorialist, it could easily supply a reason for affirming instances of Separation. When asked whether those members of a set A that have property P form a set, the NST would respond, "Of course; *any* members of A form a set!" But this is not quite the way the NST behaves. The NST will not apply the combinatorialist principle until it has assured itself that the proposed subset is definable in "an austere sublanguage of English that corresponds to the limited resources of first-order [set theory]." So the NST would seem to behave as follows. Question: "Do the singletons in A form a set?" Response: "Of course! *Any* members of A form a set and, furthermore, the property of being a singleton is first-order definable in the vocabulary of pure set theory." Now this is very strange. If *any* members of A form a set, then the P-ish members will form a set whether or not P is definable in some special way. So the reference to the definability of singletonhood adds nothing to the argument. (Well, to be fair, it does reassure us that our reasoning about singletons will not lead to any extra-mathematical paradoxes such as Richard's. But we can be confident of that as soon as we recognize that singletonhood is definable in the vocabulary of informal set theory. There is no need to insist on firstorderizability.) Consider another example: a combinatorialist can say exactly why the standard members of ω form a set without having to reflect on whether standardness can be defined in a vocabulary less expressive than the language of informal set theory. (You know the routine: since *any* members of ω form a set, the standard ones do.) That the NST behaves otherwise shows how hollow its commitment to COMBINATORIALISM really is.

We are stuck with a mystery. The NST *claims* to be a combinatorialist; yet it does not behave like one. Why does it include in the core only those instances of Separation that have passed a definabilist test? Indeed, why does it not include in the core the principle that any members of a set form a set? This combinatorialist thesis is not just a regulative principle governing axiom choice; it is a substantive claim about sets. We would like some *reason* for relegating this substantive claim to the penumbra. At one point, the core is characterized as the domain of object-talk (discourse about sets) while the penumbra is characterized as the domain of meta-talk (discourse about discourse about sets). But this is not helpful. When we say that "every collection that can be formed at a given stage, will be formed at that stage" we are talking about sets, not about set theoretic discourse. So the object-talk/meta-talk distinction only makes it more obscure why the combinatorialist thesis is absent from the core. We are elsewhere told that, "*V* [the universe of sets], in the penumbra, is as described by our accepted theory of sets, the theory of core" (p. 215). This suggests that the penumbra includes no substantive claims about sets not already included in the core; and this would mean that the substantive, combinatorialist thesis can appear in the NST's world only if it appears in the core—hardly a reason for including it in the penumbra while excluding it from the core. Perhaps the best way to make sense of the most text here is to treat the NST as a *very* odd sort of combinatorialist. The NST's "combinatorialism" consists in its assertion of certain meta-claims (claims about set theoretic discourse) none of which imply the characteristic combinatorialist thesis that any members of a set form a set. Instead, the meta-claims lead the NST to apply a definability test to the properties that are to figure in core instances of Separation and to the functions that are to figure in core instances of Replacement. A relatively unimportant question here is how it could possibly be helpful to use the term 'combinatorialism' to describe this feature of the NST model. More crucial is the question of why we should entertain a model of set theoretic activity that excludes the genuine combinatorialist's substantive claims about sets. I would like to consider three possible responses to this latter question.

(1) "Zermelo argued for a second-order account, but Skolem's [first-order] position carried the day." (p. 50) This could mean: substantive claims about sets made by real honest-to-gosh set theorists are all (or nearly all or mostly) firstorderizable; so modest philosophers should reflect this feature of real set theoretic practice in their model of that practice.

Well, if George Boolos was a real honest-to-gosh set theorist, then at least one real honest-to-gosh set theorist (other than Zermelo himself!) went to some trouble to formulate and endorse a version of Separation that is demonstrably nonfirstorderizable. Now it may be that the Boolos approach is eccentric (so that his style of set theoretic discourse can be excluded from our model without much compromising our model's fidelity to real life set theorizing). I can only report that I have yet to see good evidence of its eccentricity or that of any other nonfirstorderizable version of combinatorialism. It will not do simply to observe that first-order theories are the preferred objects of set theoretic metatheorizing. For it does not follow that those first-order theories best express the metatheorizers' conception of set. Furthermore, even if we could survey *every* utterance of *every* set theorist, it does not seem likely that we would come away with a clear case for a first-order reconstruction of those utterances. For example, could anyone really demonstrate that Bernays meant to express a firstorderizable claim when he remarked that a combinatorialist "views a set of integers as the result of infinitely many independent acts of deciding for each number whether it should be included or excluded"? Be that as it may, response 1 fails for lack of evidence (whether or not *we* manage to show that such evidence is unobtainable).

(Historical note: Fraenkel had to invent a firstorderizable version of set theory in order to prove the independence of Choice. A standard explanation for this is that Zermelo's notion of "definite property" was too vague to be metamathematically tractable. No doubt this is correct. But if Zermelo had managed in 1908 to express precisely the full combinatorial conception of set, then Fraenkel would still have needed to concoct an alternative theory. He could hardly have proved the independence of Choice from second-order Z. So, even at birth, first-order set theories were attractive, at least in part, because they allowed a gifted mathematician to show off a clever, interesting, and suggestive metatheorem. It does not follow that they offered the analysis or reconstruction of Zermelo's original *Aussonderung* axiom that best captures the combinatorial conception of set.)

(2) "A good tree cannot bring forth evil fruit, neither can a corrupt tree bring forth good fruit." The purpose of the NST model is to enhance our understanding of theory choice in mathematics. If it does so, it requires no further justification.

Fair enough. But does it really do so? Once upon a time, Nagel and Newman could observe with complete confidence how "clear" it is that "the proper business of the pure mathematician is to *derive theorems from postulated assumptions*, . . . it is not his concern as a mathematician to decide whether the axioms he assumes are actually true." That is, "the sole question confronting the pure mathematician . . . is not whether the postulates he assumes or the conclusions he deduces from them are true, but whether the alleged conclusions are in fact the *necessary logical consequences* of the initial assumptions." Well, it can be a great contribution to the species for a philosopher to leave us less certain of the clarity of some "clear truths." (How fine it is that we humans have cut down a bit on our unreflective references to *the* scientific method. Thank you Paul Feyerabend!) Professor Maddy's work over the last dozen years has helped us to see how implausible the Nagel and Newman model of mathematical practice really is. (That, in itself, is a worthy legacy for a philosopher.) Mathematicians do choose axioms. They *must* choose axioms. It seems arbitrary to declare this activity extra-mathematical. (Indeed, it seems *wrong* given that folks like Zermelo, Skolem, and von Neumann owe part of their *mathematical* reputation to their skill as axiomatizers.) Furthermore, it is, to say the least, immodest for the philosophers (the *philosophers* mind you) to *assume* that these choices are nonrational. We should at least take a look. We should at least *try* to detect rational grounds for theory choice in mathematics. The NST model is meant to advance this worthy project. The question, again, is whether it really does so.

Maddy's test case is Gödel's axiom of constructibility. There is "deep and widespread resistance to adding $V = L$ as a new axiom" (p. 129). The NST model is supposed to help us understand why this resistance is rational. Furthermore, set theorists advance reasons for their resistance: "$V = L$ is restrictive, limiting, minimal, and . . . these things are antithetical to the general notion of set" (p. 84). One would like the NST model to help us understand these reasons.

L is, of course, minimal in a very strong sense. There is a theorem θ of $ZF - P$ (ZF minus Power Set) with the remarkable property that every transitive proper class ϵ-model of θ contains L. So an ϵ-model of a certain little bit of ZFC will leave out a constructible set only if it leaves out a member of one of its own members or, if transitive, leaves out most ordinals. So every plausible candidate for V will contain L. (So L is the least plausible candidate. Does that mean it is the least plausible?) If V *can* be any thicker than L (if $ZF^2 + V \neq L$ is consistent), then, according to combinatorialist doctrine, V *is* thicker than L and, hence, *every* transitive proper class ϵ-model of θ distinct from L will approximate V better than L. If V *cannot* be any thicker than L (if $ZF^2 + V \neq L$ is inconsistent), then every transitive proper class ϵ-model of θ will approximate V exactly as well as L (because it will *be* L). In this last case, L would be both minimal and maximal and, hence, its minimality would not be a mark against it. Well, the NST appears unwilling or unable even to entertain the question of whether $ZF^2 + V \neq L$ is consistent and so must argue that L's minimality is (at the very least) unattractive without even considering whether L is, in fact, maximal.

The NST urges us to consider the following facts. We have good evidence that there are consistent extensions of ZFC that prove $V \neq L$. Let T be any one of them. (In a bit, we'll see what happens when we just let T be $ZFC + V \neq L$.) Then T proves that $ZF + V = L$ has a transitive proper class ϵ-model, but no consistent extension of $ZF + V = L$ proves that T has a transitive proper class ϵ-model. The NST concludes that $ZF + V = L$ is *restrictive* in a mathematically undesirable way. Why? We might say: while people living in a universe thicker than L can still identify the constructible sets, and so can still investigate what things would be like if V *were* L, people living in L cannot investigate what things would be like if there were nonconstructible sets. So people living with nonconstructible sets can investigate a broader range of mathematical possibilities than people living in L. That may seem plausible, but, in fact, the NST seems not to be in a position to make this argument stick.

Let us state more carefully what inhabitants of L can and cannot do. The L-ites cannot investigate what it would be like to interpret 'ϵ' as ϵ, keep every ordinal, keep every member of everything they keep, keep enough stuff to verify all of ZF, while deleting everything that allows them to tell that $V = L$. Nonetheless, if the L-ites could convince themselves that some *set* is a transitive ϵ-model of ZF, they could develop a rich account of what life would be like if nonconstructibles roamed the universe. Granted, they could imagine themselves in a nonconstructible's habitat only if they imagined away most ordinals. But, then, people who live with a measurable cardinal can imagine themselves inside L only if they imagine away most subsets of ω. So an L-model of $V \neq L$ is not to be condemned *simply* because it omits lots of stuff. Devotees of $ZFC + V \neq L$ could protest that, whereas their model of $V = L$ just *is* L, the L-models of $V \neq L$ are nothing more than countable sets, and so hardly faithful representations of their own ample universe. But why should we believe that it is even possible for a transitive ϵ-model of ZFC $+$ V $\neq$ L to be a proper class? Any evidence for the consistency of ZF^2$+$ V $\neq$ L would be evidence for the possibility of such a model. But the NST declines to consider the consistency of second-order theories and so seems poorly situated to offer any argument for the desired conclusion. It looks like a standoff. Supporters of V $=$ L can insist that all transitive ϵ-models of ZFC $+$ V $\neq$ L are sets. Supporters of V $\neq$ L can always deny this. The NST seems unable to choose rationally between these competitors.

To its credit, the NST admits as much: if the only competitor of ZF $+$ V $=$ L were "a non-starter like ZFC $+$ V $\neq$ L" it would then appear that ZF $+$ V $=$ L "is not restricting set theory from developing in any direction we might be inclined to take" (p. 230). Our comparison of ZF $+$ V $=$ L and ZFC $+$ V $\neq$ L does *not* supply us with "clear grounds for claiming that ZFC $+$ V $=$ L is restrictive" (p. 224). The NST argues instead that the restrictiveness of ZF $+$ V $=$ L emerges from a comparison with ZFC $+$ MC. The argument has two parts. First, while people living with measurable cardinals can still identify the constructible sets and so can still investigate what things would be like if V *were* L, people living in L cannot investigate what things would be like if there were measurable cardinals. (This raises a question: if this is an essential part of the argument, why was it reasonable for set theorists to believe V $\neq$ L *before* Scott proved that ZF $+$ V $=$ L $+$ MC is inconsistent?) Second, the assumption that there are measurable cardinals is mathematically attractive. So if we were all required to assume all the time that V $=$ L, the range of mathematical possibilities we are allowed to explore would be narrowed in a mathematically indefensible way. An axiom is inappropriately restrictive if it "restricts set theory from developing in a direction that has identifiable attractions" (p. 230).

Well, indeed, it would be A VERY BAD THING if a mathematics czar prevented us from deriving theorems incompatible with the axiom of constructibility. Our current knowledge of sets suggests that this would be EVEN WORSE than if the czar required our theorems to be compatible with MC. (Closet fans of constructibility could always relativize their theorems to L.) We can all agree that unreflective or involuntary allegiance to V $=$ L would have a poisonous effect on mathematical inquiry. It seems not to follow, though, that the axiom of constructibility is "antithetical to the general notion of set" (p. 84). So the NST model has not helped us to understand why it is *reasonable* to regard the axiom as antithetical. (The annihilation of the human species would have an even more unfortunate effect on mathematical inquiry; but that doesn't mean that our extinction is antithetical to the general notion of set.) Furthermore, the NST seems to employ an arbitrarily constricted standard when it argues for the attractiveness of ZFC $+$ MC. Granted, "we can entertain no reasonable hope of giving a formal criterion for 'attractiveness' " (p. 231). Yet one still craves an explanation when relevant questions are ignored. A genuine combinatorialist would regard evidence for the inconsistency of ZF^2$+$ MC as evidence for the unattractiveness of ZFC $+$ MC. Why should we believe that a creature blind to this issue is in a favorable position to assess the attractiveness of ZFC $+$ MC? Why should *we* accept the NST's judgments about attractiveness when the NST has not even entertained the question of second-order consistency? There might be a good argument for why we should, but we really need to *see* that argument.

There is one point at which the NST's narrowness of vision clearly does not serve it well: it is unable to express any convincing reservations about the odd theory ZFC + ∼Con(ZFC). Here is the NST's best effort.

> From Gödel's incompleteness theorem, we know that a theory like ZFC + ∼Con(ZFC) is consistent if ZFC itself is consistent, but given . . . our legitimate preference for consistent theories . . . and given that a metamathematical theorem like ∼Con(T) provides good evidence for the inconsistency of T, it is hard to see how ZFC + ∼Con(ZFC) could ever present itself as a serious candidate for adoption. (pp. 229–30)

Unfortunately, one of the NST's premises is false. A key lesson of Gödel's second incompleteness theorem is that, in a first-order setting, an *object language* theorem like ∼Con(ZFC) does *not* provide good evidence for the inconsistency of ZFC. The existence of a nonstandard ZFC "proof" of absurdity tells us nothing about the consistency of ZFC. Yet this is one (and probably the only) way for a first-order model to verify ∼Con(ZFC). So the NST's argument is flawed. Of course, the conclusion might still be correct. Indeed, it *is* "hard to see how ZFC + ∼Con(ZFC) could ever present itself as a serious candidate for adoption." But that's because ZF^2+ ∼Con(ZFC) is inconsistent and, hence, ZFC + ∼Con(ZFC) can have no models that agree with the full combinatorial concept of set. The NST seems unable even to express the central combinatorialist doctrine. So *this* argument is unavailable to the NST and, in its absence, the NST might indeed be tempted by ∼Con(ZFC). Why shouldn't the NST some day be convinced that ∼Con(ZFC) is a valuable postulate in the theory of nonstandard ZFC proofs? *We* could recognize then that the NST's natural numbers do not form a structure of order type ω. But why should the NST care about agreement with a structure that, from its perspective, cannot be characterized categorically? If the NST insists on exploring only one set theory at a time, then its adoption of ∼Con(ZFC) would prevent its exploration of ZFC + Con(ZFC). But why should the NST consider this a serious problem? Why would it want to explore a world in which a particular sort of nonstandard "proof" fails to exist? The assumption that such objects do exist seems more likely to be stimulating and fruitful.

To those who find this whole scenario fantastic and bizarre, I can only say: I agree!! It would be fantastic and bizarre for a typical real-life set theorist to endorse ZFC + ∼Con(ZFC). But that is because real-life set theorists typically employ (implicitly or explicitly) the full combinatorial concept of set. If the NST is unable to grasp this concept, then, as far as I can tell, it has no compelling reason to shun ∼Con(ZFC) and might be able to devise reasons for embracing it. So we should not assume that the NST is well situated to determine which set theories are attractive. And one has to question whether the NST model will really help us to understand why it is reasonable for real-life set theorists to resist the axiom of constructibility.

(3) " . . . these notions go beyond the methodologically relevant into extra-mathematical theorizing" (p. 215). If the consistency of ZF^2+ MC were some sort of metaphysical pseudoproblem, then it would not be "methodologically relevant" to the question of whether ZFC + MC is mathematically attractive.

Well, I'm sure of this much: we do not currently *know* that the consistency of ZF^2+ MC is a metaphysical pseudoproblem. So it would not be reasonable to dismiss it as such. (A question: which sort of statement admits of the more convincing evidence; “Such and such is a metaphysical pseudoproblem” or “So and so is a consistent second-order theory”?) It’s not as if questions of second-order consistency were *in general* mathematically intractable. We have excellent evidence, for example, that Z^2 is consistent: Z^2 would be true in $V(\omega+\omega)$, if there were such a thing as $V(\omega+\omega)$; furthermore, the experience of the hundreds of massively intelligent people who have fiddled with the relevant concepts suggests strongly that our notion of $V(\omega+\omega)$ is not fundamentally incoherent; so we have good reason to believe that $V(\omega+\omega)$ is logically possible; so we have good reason to believe that Z^2 is consistent. But how might we show that ZF^2+ MC is consistent? I have nothing too helpful to offer on that front. The best I can do is to indicate why I do not claim to know that the project is incoherent or unintelligible.

3 How do we know what’s possible? If, necessarily, objective empirical knowledge is possible only if every event has a cause and if, furthermore, objective empirical knowledge *is* possible, then it is a necessary truth that every event has a cause. This may appear to be a promising technique for deriving the necessity of determinism from the bare possibility of objective empirical knowledge: we need not convince the sceptic that we do possess objective empirical knowledge; we need only show that we *could* possess it. But how do we show that we could without first showing that we do? How do we show that something is possible without showing that it is actual?

Well, even though I’m not able to manage a transcendental argument for determinism, the general sort of inference that’s required is perfectly common. Here is an example. The *tee game* is played on a board with 15 holes arranged in rows of 1, 2, 3, 4, and 5 to form an equilateral triangle. The game begins with all but one hole filled with a playing piece (typically a golf tee). The player removes pieces by jumping (as in checkers). The goal is to jump until just one piece is left. (It’s easy to get caught with more than one piece and no possible jumps.) Label the holes as follows. Call the three vertices of the triangle the ‘A’ holes. Call the six holes adjacent to the A holes on the outer edge the ‘B’ holes. Call the three holes in the middle of an outer edge ‘C’ and call the three interior holes ‘D’. We can now characterize a game in terms of its initial and final positions. For example, an AC game would begin with an A hole empty and end with the one remaining piece in a C hole. It is now natural to ask which of the 16 types of game are possible. How might we show that, say, a CD game is possible? Well, we could *play* a CD game by moving pieces around the board; but that’s not required. Here’s one reason why.

Every sequence of tee game moves has a *dual* sequence formed by interchanging full and empty holes and running the sequence backward. So, for example, a game that begins with an empty D and ends with a filled C has a dual game that ends with a filled D and begins with an empty C. So we could show that a CD game is possible by actually playing a DC game. In fact, I *did* infer the possibility of a CD game from my friend Bob Graber’s having played a DC. At that point, I knew that it would not be fundamentally incoherent for someone to claim to have played a CD game; and when Bob returned from vacation to announce triumphantly that he had managed to do so, I didn’t doubt him for an instant.

How should we understand this? There seems no mathematically significant difference between a given tee game and its dual. So the characteristically mathematical move would be to regard the physical performance of a DC game and the physical performance of a CD game as instantiations of the same abstract structure (abstract, at least in part, because in contemplating the structure we abstract from those features that allow us to distinguish between a game and its dual). The physical performance of a DC game shows a CD game to be logically possible because it *itself* possesses all the mathematically significant features of a CD game: the DC game instantiates the DC/CD structure. The physical performance of a DC game does not in itself establish the *physical* possibility of a CD game. (I mean the physical possibility of the physical performance of a tee game in which the initial empty hole is a C and the final occupied hole is a D—where 'occupied', 'empty', 'initial', and 'final' are understood in the natural way.) We have no *mathematical* guarantee that a CD game is permitted by physical law. Our confidence that it is permitted is a product of *physical*, not just mathematical, insight. Maybe Graber lied and strange physical forces *do* prevent the completion of a CD game! The game is still logically possible and we still *know* it to be logically possible.

The point is that we possess a notion of logical possibility distinct from physical possibility and we really do manage to acquire knowledge about what is logically possible; furthermore, one technique for doing so is to show that superficially distinct structures are, from some mathematically legitimate perspective, indistinguishable. Perhaps, a thousand years from now, someone will notice that, viewed in the right way, a model of $ZF^2 +$ MC is indistinguishable from a structure whose logical possibility is well established. Or perhaps such an idea is fanciful. I can only report that I am not nearly sophisticated enough to be confident that such a thing will *not* happen. A more general question: could a model of $ZF^2 +$ MC result from some possibility-preserving transformation of a structure reasonably thought to be possible? We have examples of such arguments: the consistency of second-order arithmetic follows from the possibility of a model for Robinson arithmetic, since a model of the former can be extracted from a model of the latter. The trick is to use the second-order machinery to expel alien intruders. I have no idea whether a model of $ZF^2 +$ MC could be extracted from a structure that is both richer and more easily seen to be coherent. Again, I only report that I lack the sophistication to establish that this will never happen.

There is something else worth noting about the tee game example. One way of establishing that a situation is logically possible is to describe a set theoretic model of a formal theory that expresses the mathematically significant features of that situation. Evidently, though, this is not the only way to establish logical possibility, since this is not how we established the possibility of a CD game. I have no doubt that we could produce a model theoretic argument. But our duality argument is not just a way of convincing ourselves that a model theoretic argument can be concocted. It's not as if the only genuine possibility arguments are model theoretic, whereas other techniques serve only to confirm that those arguments are available. The duality argument convinces us that a model theoretic argument is available because we know model theory to be a successful regimentation of preexisting techniques for establishing logical possibility. It would be astounding if the duality argument involved a technique unanticipated by the model theorists (especially since the duality reasoning is entirely finitistic). If, contrary to expectation, a model theoretic reconstruction eluded us, we would want to look more closely at the duality argument—but we would also have to consider augmenting model theory! Model theory does not legislate the notion of logical possibility. It is model theory that is expected to accommodate the full range of unregimented, but nonetheless legitimate, conceptions of logical possibility. It is these preexisting conceptions that supply the standard by which model theory is to be judged.

The moral is: though our concept of logical possibility may not live up to popular notions of mathematical precision, it is not extra-mathematical or methodologically irrelevant. Now I admit, it does not follow that this notion is sufficiently determinate to fix a truth value for every claim about second-order consistency. But it *might* be sufficiently determinate to fix a truth value for the claim that $ZF^2 + MC$ is consistent. And even if it does not do that in its current form, there might be some mathematically natural precisification of the notion that allows it to fix a truth value. We will not know one way or the other unless we do a good bit of fiddling with second-order theories.

4 How do we know what's impossible? Possibility is to consistency as impossibility is to consequence. Let us consider an informal impossibility argument. The six types of tee game move are listed in the table below along with the effect of each move on the number of pieces occupying the four types of holes. The table also tracks the *sum* of the A and C pieces.

	ΔA	ΔB	ΔC	ΔD	$\Delta A + \Delta C$
ABC	-1	-1	$+1$	0	0
BCB	0	0	-1	0	-1
BDD	0	-1	0	0	0
CBA	$+1$	-1	-1	0	0
CDC	0	0	0	-1	0
DDB	0	$+1$	0	-2	0

An example may help to make clear how the table is to be understood. Consider line 6: DDB. A DDB move is one in which a piece occupying a D hole jumps an adjacent piece occupying a D hole and moves into an empty B hole. The jumped piece is removed, while the jumping piece shifts from a D to a B. The effect is that there is one more occupied B hole ($+1$ in column 2) and there are two fewer occupied D holes (-2 in column 4).

The table makes three important facts evident: there is no way to increase the number of occupied D holes; there is only one way (DDB) to increase the number of occupied B holes; there is only one way (BCB) to decrease the sum of occupied As and Cs.

Now consider the DD game (initial empty position = D; final occupied position = D). We are to begin with two occupied Ds and end with one. Since Ds cannot be restored, we can lose only one in the course of the game. This means we are allowed exactly one CDC and no DDBs. (A DDB would wipe out all our Ds.) Since we begin with three As and three Cs, we require six BCBs to eradicate them. With no DDBs, there is no way to increase the number of Bs—either overall or on a given edge. Focus on just one edge and suppose we have just performed a BCB there. We now have exactly one B and no Cs. We need to restore the C if we are to perform another BCB.

If we restore the C with an ABC, we will have permanently lost our Bs and no further BCBs will be possible. So we must use up our one CDC to restore the C. At most one BCB will then be possible on each of the two remaining edges. So we can perform at most four BCBs. So the DD game is impossible. (There are, in fact, five types of impossible game: AD, BD, DA, DB, and DD.)

I am confident that we could construct a first-order theory of the tee game one of whose theorems is that no DD game is possible. But our informal impossibility argument is not just a way of convincing ourselves that we can manufacture a first-order derivation. We are confident that a first-order characterization of the tee game is available because we know first-order languages to be successful regimentations of preexisting forms of finitistic expression. Our informal argument convinces us that a first-order derivation is obtainable because we know first-order logic to be a successful regimentation of preexisting techniques for establishing logical impossibility (in particular, the logical impossibility of certain premises being true while a certain conclusion is false). Our notion of logical impossibility supplies a standard by which formal logics are to be judged. So this notion is not extra-mathematical or methodologically irrelevant. It does not follow that our conception of logical impossibility or fundamental incoherence is sufficiently determinate to fix a truth value for every claim about second-order consequence. But we will have no hope of discovering the limitations of the concept if we refuse to think about it.

5 Overall assessment Professor Maddy has crafted an important and promising research program. Mathematicians must choose axioms. We philosophers should not assume that they behave nonrationally when they do so. Instead, we should try "to explicate this practice in such a way as to lay bare its underlying rationality" (p. 233). Maddy has positioned herself well to carry through this program. She does not claim, though, to have progressed far.

> My hope is that the beginnings sketched here are compelling enough to inspire those cleverer and more knowledgeable than myself—to correct my errors, to fill in what's been passed over in the case against $V = L$, and to extend naturalistic methods to the evaluation of higher and more controversial hypotheses. (p. 234)

While I must agree that the main contribution of the present book is to set the stage for future progress, I am not persuaded by Professor Maddy's explanation for the preliminary character of her work. Even clever and knowledgeable people can be hobbled by philosophical prejudice—and it is philosophical prejudice, not insufficient brain power, that seems to impede Professor Maddy. Maddy may not be positively convinced that the consistency of, say, $ZF^2 + MC$ is a metaphysical pseudoproblem, but, evidently, she leans far enough in that direction to feel comfortable about passing over the issue in silence. This will not do. It turns out that the antics of a firstorderized homunculus do little to illuminate axiom choice in set theory. So we must face the question of which nonfirstorderizable locutions should be introduced to the NST and what it should do with them. A mastery of plain old plural quantifiers would allow the NST to become a genuine combinatorialist. (And, like every real mathematician, the NST could then understand what ω looks like!) If this enhancement of the NST is offensive to philosophical principles, then mathematics itself must offend them. Modesty requires that such principles be jettisoned.

Philosophy will never be as much fun as mathematics; but it might help us to think about the mathematics more clearly. If it does the opposite, good riddance!

Stephen Pollard
Division of Social Science
Truman State University
Kirksville, Missouri 63501
email: spollard@truman.edu