Neutralism within the semantic tradition

Robert Trueman

This is my final draft of an article forthcoming in *Thought*

1. Introduction

Neutralism is the thesis that second-order quantification cannot generate an ontological commitment to a kind of entity that is not already generated by the use of predicates.[1] In many ways, neutralism is an obvious stance to take, but it has never reached the status of orthodoxy. This is due in large part to the force of Quine's polemics against second-order logic: Quine (1970: 66–8) famously thought that predication comes at no ontological cost but second-order logic is "set-theory in sheep's clothing".

In this note I will argue for the viability of neutralism. Importantly, in doing so I want to remain steadfastly agnostic about what the ontological commitments of second-order quantification actually are. What I will do, then, is try to present what I will call a *neutralist framework*: an account of the second-order quantifiers which does not by itself tell us what the commitments of second-order quantification are, but which does tell us that these commitments cannot exceed those of predication. Wright (2007) has recently argued that an inferentialist account of the second-order quantifiers would serve as an adequate neutralist framework. Now, I have no desire to quibble with this claim here.[2] What I do want to show, however, is that we do not *have* to become inferentialists in the pursuit of a neutralist framework. Nor, for that matter, do we have to become substitutionalists.[3] Rather, a neutralist framework can be established within the mainstream semantic tradition.

Before going any further, I must address one possible confusion. In the next section I will offer a semantics for the second-order quantifiers, and in doing so I will assign sets to predicates. This may lead the reader to think that on this semantics, it is just obvious that predicates refer to sets. But the validity of this inference will, of course, depend on what we mean by 'refer'. Throughout this paper I will

[1] This definition of 'neutralism' is essentially Wright's (2007: 153). However, whereas Wright takes neutralism to be a thesis concerning quantifiers of all kinds, I have restricted its scope to second-order quantifiers alone. This restriction is merely for ease of expression.

[2] However, in §4 I will briefly contrast Wright's neutralist framework with my own.

[3] See (Barcan-Marcus 1972: §III) for a discussion of the relation between neutralism and substitutional treatments of the quantifiers.

take it that what it means to say that a given expression refers to a given entity is that any sentence featuring that expression *says something of* that entity. When we understand 'refers' in this way, the mere fact that we assign sets to predicates in the process of formulating a semantics does not entail that predicates refer to sets on that semantics. Indeed, this is a point on which I believe Quine and I agree. Although he wanted to deny that predicates refer to sets, he saw that he was not thereby forced to

> deny that there are certain sets connected with [predicates] otherwise than in the fashion of being referred to. On the contrary, in that part of the theory of reference which has to do with sets there is occasion to speak of the *extension of* a general term or predicate — the set of all things of which the predicate is true. One such occasion arises when in the theory of reference we treat the topic of validity of schemata of pure quantification theory [...] The general theory of quantificational validity thus appeals to classes, but the individual statements represented by the schemata of quantification theory need not; the statement '$(\exists x)(x$ is a dog . x is black)' involves, of itself, no appeal to the abstract extension of a predicate. (Quine 1951: 95)[4]

2. A semantics for the second-order quantifiers

In this section I will outline a formal semantics for a second-order language. This semantics is an extension of the β-variant semantics for a first-order language.[5] I am not the first to give this semantics; Boolos presented it in his (1975: 513–4), before he hit upon his plural treatment. In the next section I will argue that this semantics is a neutralist framework. To be clear, I have chosen this semantics over the more familiar Tarskian alternative not for philosophical reasons but for presentational ones: I think it is simply easier to see the relation between neutralism and my preferred semantics.

Let $\mathcal{L}$ be a formal second-order language of the usual kind. Rather than begin with $\mathcal{L}$ in its entirety, it will be useful to start by giving a semantics for its first-order fragment. As is usual, an interpretation is any ordered pair, $\langle D, v \rangle$, where D is a domain of objects and v is a valuation function; v maps individual constants from $\mathcal{L}$ to objects and n-adic predicates of $\mathcal{L}$ to sets of ordered n-tuples of D.[6]

[4] For the sake of conformity with my own terminology, I have replaced every occurrence of 'class' in the above quotation with 'set', and one occurrence of 'named' with 'referred to'; I do not believe that this last alteration has done any violence to Quine's meaning. A near identical passage appears in (Quine 1980: 115). See also (Quine 1970: 67).

[5] See (Mates 1972: ch.4) for a detailed exposition of this semantics.

[6] As I will mention in §3, the choice to assign sets to predicates rather than, say, properties, is inconsequential.

Let $\mathcal{I}$ be any interpretation, Φ, Ψ and Ξ be any wffs of $\mathcal{L}$, α be any first-order variable of $\mathcal{L}$, and β be any individual constant of $\mathcal{L}$ that does not appear in Φ

(1) If Φ is an n-adic predicate letter followed by n individual constants, then Φ is true on $\mathcal{I}$ iff the n-tuple of objects that $\mathcal{I}$ assigns to the individual constants in Φ, taken in the order in which those constants occur, is a member of the set that $\mathcal{I}$ assigns to the predicate in Φ

(2) If $\Phi = \ulcorner\neg\Psi\urcorner$, then Φ is true on $\mathcal{I}$ iff it is not the case that Ψ is true on $\mathcal{I}$

(3) If $\Phi = \ulcorner\Psi \wedge \Xi\urcorner$, then Φ is true on $\mathcal{I}$ iff Ψ is true on $\mathcal{I}$ and Ξ is true on $\mathcal{I}$

(4) If $\Phi = \ulcorner\exists\alpha\Psi\urcorner$, then Φ is true on $\mathcal{I}$ iff $\Psi[\alpha/\beta]$ is true on some β-variant of $\mathcal{I}$

(5) If $\Phi = \ulcorner\forall\alpha\Psi\urcorner$, then Φ is true on $\mathcal{I}$ iff $\Psi[\alpha/\beta]$ is true on every β-variant of $\mathcal{I}$

$\Psi[\alpha/\beta]$ is the result of replacing every occurrence of α in Ψ with β; for example, 'x is wise'['x'/'Socrates'] is 'Socrates is wise'. A β-variant of an interpretation $\mathcal{I}$ is an interpretation that differs from $\mathcal{I}$ only by assigning a different object to β, if it differs at all. This semantics for the first-order fragment of $\mathcal{L}$ has a passing similarity with a substitutional semantics, but should not be confused with it. On the β-variant account, '$\exists x\, x$ is wise' can be true on an interpretation $\mathcal{I}$ even if no instance of 'x is wise' is true on $\mathcal{I}$: all that is required is that $\ulcorner\beta$ is wise$\urcorner$ be true on some β-variant of $\mathcal{I}$.

This semantics is straightforwardly expanded to include the second-order quantifiers. All we need to do is add the following clauses:

Let A be any n-adic second-order variable of $\mathcal{L}$, and B be any n-adic predicate of $\mathcal{L}$ that does not appear in Φ[7]

(6) If $\Phi = \ulcorner\exists A\Psi\urcorner$ then Φ is true on $\mathcal{I}$ iff $\Psi[A/B]$ is true on some B-variant of $\mathcal{I}$

(7) If $\Phi = \ulcorner\forall A\Psi\urcorner$ then Φ is true on $\mathcal{I}$ iff $\Psi[A/B]$ is true on every B-variant of $\mathcal{I}$

A B-variant of $\mathcal{I}$ is, of course, an interpretation that differs from $\mathcal{I}$ only by assigning a different set of ordered n-tuples to B, if it differs at all. Again, this semantics for the second-order quantifiers has a passing similarity with a substitutional semantics, but again it should not be confused with a substitutional semantics: on the B-variant account, '$\exists F\, F$(Socrates)' can be true on an interpretation $\mathcal{I}$ even if no instance of 'F(Socrates)' is true on $\mathcal{I}$.[8]

[7] If any predicates are being treated as logical constants, for example '$x = y$', then B must be any predicate that is not a logical constant.

[8] This semantics is also similar in some respects to a heuristic that Wright (2007: 161–2) offers

3. A neutralist framework

There is a clear affinity between the B-variant semantics that I have just outlined and neutralist frameworks. When we expanded the β-variant semantics for the first-order fragment of $\mathcal{L}$ to include the second-order sentences, we did not add extra domains of entities to our interpretations for the second-order quantifiers to quantify over; nor did we import any extra members into the first-order domains. All we did was add clauses telling us which of the second-order sentences are true on which of the interpretations that we already recognised. '$\exists F\, F$(Socrates)', for example, is true on interpretation $\mathcal{I}$ iff there is some B-variant of $\mathcal{I}$ on which $\ulcorner B$(Socrates)$\urcorner$ is true. Whatever ontological force the second-order quantifier in '$\exists F\, F$(Socrates)' has must therefore be traced back to the way these B-variants interpret B. But B is a perfectly ordinary predicate, and each B-variant interprets B in the same way that it interprets every other predicate.

We can travel further down this line of thought by asking whether or not second-order quantifiers quantify over sets on the B-variant semantics. In doing so we must be careful not to confuse the questions: (i) 'On the B-variant semantics, do we quantify over sets when we use the second-order sentences of $\mathcal{L}$?' and (ii) 'Do we quantify over sets when we give the B-variant semantics for $\mathcal{L}$?' The answer to (ii) is obviously, 'Yes, the B-variant semantics is set-theoretic through and through!' But it is with (i) that we are really interested, and (i) and (ii) are no more equivalent than 'Do we quantify over linguistic formulae when we use a given language' and 'Do we quantify over linguistic formulae when we give a semantics for that language?' As we will see, how we should answer (i) depends wholly on what we think predicates do; in particular, on whether we think that they refer to sets.

Suppose first that we do think that predicates refer to sets; that is, we think that on any given interpretation, 'Socrates is wise', for example, says something of the set assigned to 'x is wise'. In that case, we absolutely should say that second-order quantifiers quantify over sets, just as we say that *first-order* quantifiers quantify over *objects*. To say that first-order quantifiers quantify over objects on the β-variant semantics is to say that a certain relation holds between using a first-order quantifier and saying something of an object: '$\exists x\, x$ is wise' is true on $\mathcal{I}$ iff $\ulcorner\beta$ is wise$\urcorner$ is true on some β-variant of $\mathcal{I}$, and on each β-variant $\ulcorner\beta$ is wise$\urcorner$ says something of the object referred to by β. If predicates refer to sets, then an exactly analogous relation holds between using a second-order quantifier and saying something of a set: '$\exists F\, F$(Socrates)' is true on $\mathcal{I}$ iff $\ulcorner B$(Socrates)$\urcorner$ is true on some B-variant of $\mathcal{I}$, and on each B-variant $\ulcorner B$(Socrates)$\urcorner$ says something of the set referred to by B.

as an aid to understanding his own inferentialist account. The most important difference is that Wright's heuristic makes dangerous play with propositions (which is why Wright treats it as a mere heuristic), and this semantics does not.

Now suppose that we do not think that predicates refer to sets. As I mentioned in §1, this will not compel us to abandon the B-variant semantics. Nor, for that matter, will it force us to relegate that semantics to the status of a mere mathematical tool for defining logical consequence.[9] All we have to do is follow Quine's lead and insist that interpretations do not assign sets to predicates as their referents, but as their extensions. When we understand our interpretations in this way, we will deny that we say something of the set assigned to a predicate by using that predicate. And in that case it is a mistake, or at least badly misleading, to say that second-order quantifiers quantify over sets.[10] It is a mistake because it invites us to draw the analogy just discussed with the claim that first-order quantifiers quantify over objects. If predicates do not refer to sets then this analogy does not hold. '$\exists F\, F$(Socrates)' is true on $\mathcal{I}$ iff $\ulcorner B$(Socrates)$\urcorner$ is true on some B-variant of $\mathcal{I}$, but on no B-variant does $\ulcorner B$(Socrates)$\urcorner$ say something of the set assigned to B. In short, if predicates do not refer to sets then on the B-variant semantics, it is a pun on 'quantify over' to say that first-order quantifiers quantify over objects and second-order quantifiers quantify over sets.

So, if we intend the claim that second-order quantifiers quantify over sets to mean, roughly, that we can use second-order quantifiers to talk about sets, then on the B-variant semantics, that claim is appropriate only if we can already use predicates to talk about sets. And this, it seems to me, is just an informal way of saying: on the B-variant semantics, the use of second-order quantifiers commits us to the existence of sets only if the use of predicates does too. Moreover, this argument obviously generalises. In presenting the B-variant semantics, I assigned sets to predicates, but we could assign them any kind of entity we liked; we are, for example, free to assign properties to our predicates, or at least we are if properties exist. (Indeed, we could even refuse to assign them any entity at all!) The point remains that whatever kind of entity we assign to predicates, second-order quantification will commit us to entities of that kind on the B-variant semantics only if predication does too. The B-variant semantics is, therefore, a neutralist framework.

4. Conclusion

Wright asks us to choose between a neutralist framework and a semantic account of the second-order quantifiers. But there is no need to make that choice. We can develop a neutralist framework without abandoning the semantic tradition. (Of course, this is hardly bad news for Wright, whose primary ambition was simply to defend neutralism.) What is more, there is some reason to prefer my semantic brand

[9] Of course, this is not to say that there might not be *other* reasons to relegate the B-variant semantics in this way.

[10] The Boolos of (1975: 511) fell into this mistake.

of neutralism over Wright's inferentialist one. As Wright (2007: 166–8) admits, the incompleteness of full second-order logic is inevitably going to prove problematic for any attempt to give an inferentialist treatment of that logic. Now, while Wright (2007: 168) does suggest one way around this problem, it is worth noting that as my neutralist framework is semantic in nature, it faces no special problem when it comes to incompleteness. This is not to say that no questions remain about the strength of the second-order logic delivered by the B-variant semantics; they now reappear as questions about the strength of the set-theory in which that semantics is embedded.[11] But *if* our set-theory supplies the full stock of sets, then the B-variant semantics will yield a neutralist treatment of full second-order logic.[12]

References

Barcan-Marcus R. 1972. Quantification and ontology. *Noûs* 6: 240–50.

Boolos G. 1975. On second-order logic. *The Journal of Philosophy* 72: 509–27.

Mates B. 1972. *Elementary Logic*. 2nd edition. Oxford: Oxford University Press.

Quine W. 1951. Semantics and abstract objects. *Proceedings of the American Academy of Arts and Sciences* 80: 90–6.

———. 1970. *Philosophy of Logic*. Englewood Cliffs, NJ: Prentice-Hall.

———. 1980. Logic and the reification of universals. In *From a Logical Point of View,* 102–29. Cambridge, MA: Harvard University Press. 2nd revised edition.

Wright C. 2007. On quantifying into predicate position: steps towards a new(tralist) perspective. In M. Leng, A. Paseau & M. Potter (eds), *Mathematical Knowledge,* 150–74. Oxford: Oxford University Press.

[11] Or whatever background theory in which we embed our semantics.

[12] Thanks to Daniel Brigham, Tim Button, Owen Griffiths, Fraser MacBride, Michael Potter, Lukas Skiba and three anonymous referees. Thanks also to the Analysis Trust for their studentship.