THE JOURNAL OF
PHILOSOPHY

THE JOURNAL OF PHILOSOPHY

＋ • ＋

THE UBIQUITOUS PROBLEM OF EMPTY NAMES*

U ntil a little under thirty years ago, the orthodox view about proper names was that they had a meaning (or sense), that this meaning could be captured by a description of some kind or other, and that these names managed to refer to a particular individual because the descriptions associated with the names were definite descriptions. Something like this view was originally and independently advocated by Russell and Frege, and championed more recently by philosophers such as John Searle and Michael Dummett.[1]

[1] Of course, this is not a complete specification of the descriptive theorist's position. She will have to say much more before she can claim to be proffering a theory about the semantics of proper names. The descriptivist will at least have to provide an answer to the following questions:

(i) What kind of association is to hold between the name and the description? Can the name-description pair vary from speaker to speaker? Are they occasion dependent, so that even for a single speaker, there may be different name-description pairs on different occasions?

(ii) How close does the fit between an individual and set of descriptions have to be before the name associated with the descriptions is to count as referring to that individual?

(iii) Must the relevant descriptions associated with a name pick out a unique individual? What happens if there is more than one candidate for the referent of the name?

(iv) How are we to analyze the relevant definite descriptions? Should we follow Frege and treat them as singular terms ("Über Sinn und Bedeutung," *Zeitschrift für Philosophie und Philosophische Kritik*, C (1892): 25–50, "On Sense and Meaning," Max Black, trans., *Philosophical Review*, LVII (1948): 209–30)? Should we follow Russell and treat them quantificationally ("On Denoting," *Mind*, XIV (1905): 479–93)? Should we follow Delia Graff and treat them (at least sometimes) as predicates ("Descriptions as Predicates," *Philosophical Studies*, CII (2001): 1–42)?

(v) What kind of description will count as a reasonable candidate for capturing the sense or meaning of a proper name? Saul Kripke has shown us that not just any uniquely identifying description will do.

Descriptive theories have recently been contrasted with direct reference theories. The most basic tenet of any direct reference theory is that the semantic value of any proper name is nothing over and above the individual to which it refers. According to this view, the meaning of a name cannot adequately be captured by a description of any kind. Direct reference theorists typically supplement this tenet with the further thesis that any simple sentence containing an occurrence of a proper name expresses a singular proposition—a proposition in which the individuals referred to occur as constituents—if it expresses a proposition at all. This view was originally advocated by Mill, has been vigorously defended by authors such as David Kaplan, Nathan Salmon, and Scott Soames, and has been supported by arguments found in the work of Keith Donnellan, Saul Kripke, Ruth Barcan Marcus, and Hilary Putnam. Indeed, it would not be unreasonable to suggest that this view has become the new orthodoxy.

When philosophers do their accounting—that is, when they come to compare the merits and shortcomings of the two theories—it is often suggested that sentences containing an occurrence of an empty name are not well accounted for by the direct reference theorist, whereas descriptive theories can handle them with ease. Such a pronouncement has been made by descriptivists and direct reference theorists alike.

Descriptivists have often taken such considerations as a significant point against—if not a complete refutation of—direct reference theories, without even conceiving that similar considerations might also tell against their particular variety of descriptivism. To illustrate, consider the following:

> Russell instructs us to test a logical and semantical theory by "its capacity for dealing with puzzles." His own theory of ordinary proper names nicely passes muster with respect to…[two] such puzzles: those presented by empty proper names, [and, as a special case] by negative existentials containing proper names…. On the other hand the anti-Fregean view… founders on these rocks.[2]

> [The Kripkean] criticism of the description theory has been widely accepted as showing that no description theory of proper names is tenable. As a consequence…the problem that Millians face in the area of language has gone virtually unnoticed. This is unfortunate. The cause of philosophical understanding as well as the cause of descriptivism would have been better served had the problem received the attention

---

[2] Alvin Plantinga, "The Boethian Compromise," *American Philosophical Quarterly*, xv (1978): 129–38, here p. 129–30.

it deserves…[for] in the area of language, a strong case can be made against Millianism, in connection with the phenomena of names without bearers.[3]

Likewise, direct reference theorists tend to acknowledge that empty names present a problem for their theory. But when they consider the problem directly, direct reference theorists either attempt to explain away the problem or point to more general issues in the philosophy of language which are supposed, on balance, to vindicate their semantic theory. The interesting point to notice when one comes across any such evaluation, however, is that it is almost universally conceded—either implicitly or explicitly—that the descriptivist has no problem on this score. Consider the following appraisal of the philosophical landscape set forth by three direct reference theorists:

> There are two major types of theories of the meaning of proper names: Millian and Fregean…. Each theory (type) has its apparent successes and failures, and, often, what is regarded as one's failure is viewed as its rival's success…. [There are] two such problems for the Millian theory…the problem of negative existentials…and the problem of non-referring names.[4]

> The issues surrounding the use of non-denoting names are of *particular* interest to those philosophers who hold to a version of the direct reference theory for names and indexicals…Russell solves the problem of non-denoting proper names by claiming that proper names should be semantically viewed as definite descriptions…. From the direct reference perspective, however, Russell's puzzle about non-denoting proper names remains a puzzle.[5]

> In "On Denoting," Russell trumpeted his Theory of Descriptions…for its ability to handle a variety of puzzles that arise on his theory that the semantic content of a singular term is solely its referent (denotation, designatum). The puzzles are primarily: Frege's Puzzle about $\lceil \alpha = \beta \rceil$; the more general problem of substitution failure in certain contexts, especially those ascribing propositional attitude; the question of content and truth value for sentences involving nonreferring terms; and as a special case, true negative existentials…. It has been objected that the

---

[3] Jerrold Katz, "Names without Bearers," *Philosophical Review*, CIII (1994): 1–39, here pp. 2–3.

[4] T.C. Ryckman, "The Millian Theory of Names and the Problem of Negative Existentials and Non-referring Names," in D.F. Austin, ed., *Philosophical Analysis: A Defense by Example* (Boston: Kluwer, 1988), pp. 241–50, here p. 241.

[5] Gregory Fitch, "Non Denoting," *Philosophical Perspectives*, VII (1993): 461–86, here pp. 461–62.

second two problems are sufficient by themselves to refute Millianism even if the first two problems are not.[6]

Salmon later essentially concedes that the problem is a special problem for direct reference theorists:

> A Millian like myself...may not avail him/herself of the Theory of Descriptions to solve the problem of sentences with non-referring names. If α is a proper name, referring or not, it is not a definite description, nor by the direct reference theorist's lights does it 'abbreviate' any definite description. Direct reference theory thus excludes application of the Theory of Descriptions [to solve this problem].... For similar reasons, the direct reference theorist is also barred from using Frege's sense-reference distinction to solve the difficulties (*ibid.*, p. 285).

It is my contention, however, that this common perspective is mistaken. To the extent that the direct reference theorist has a problem with empty names, the descriptivist has one also, and it is the aim of this paper to explain why.

I. DIRECT REFERENCE (AND THE PROBLEM OF EMPTY NAMES)

In order to illustrate the problem empty names present the direct reference theorist, consider the following sentences.

(1.1) Sherlock Holmes exists.
(1.2) Ebenezer Scrooge exists.
(1.3) Ebenezer Scrooge is a detective.
(1.4) Sherlock Holmes is a detective.
(1.5) Ebenezer Scrooge does not exist.
(1.6) According to the novel, *A Christmas Carol,* Ebenezer Scrooge exists.
(1.7) According to the Conan Doyle stories, Sherlock Holmes is a detective.

Common sense tells us that sentences (1.1)–(1.7) are meaningful— that is that they express propositions—and that (at least some of) these propositions have a classical truth value. Sentences (1.1) and (1.2), for example, express false propositions; sentences (1.5)–(1.7), on the other hand, express true propositions. It has sometimes been maintained that the direct reference theorist cannot agree (cf. Katz, *op. cit.*). The worry can be put as follows: sentences containing a (non-mentioned) occurrence of a nonreferring name obtain their meaning compositionally; their meaning is determined in some way by the meaning of their parts and the manner in which they are combined. Each part of a sentence plays some role in determining which proposi-

---

[6] Nathan Salmon, "Nonexistence," *Noûs*, XXXII (1998): 277–319, here pp. 277–78.

tion is expressed by that sentence. Thus, such a sentence expresses a proposition only if each of its parts has some meaning. But according to the direct reference theorist, the meaning or semantic value of a name just is its referent. She thus seems forced to concede that sentences containing a nonmentioned occurrence of an empty name express no proposition at all. The embarrassment for the direct reference theorist, then, is that sentences (1.1)–(1.7) do seem to express propositions.

It has often been noted, however, that the direct reference theorist can provide a simple and straightforward response to this kind of objection. She might stipulate that a simple subject/predicate sentence will be true if and only if the subject expression refers to an object (or objects) which has (have) the property expressed by the predicate; it will be false otherwise.[7] Thus, any sentence of the form 'a is F' will be false if the name 'a' has no referent. It will also be part of this view that when such sentences occur within the scope of a sentential operator, the complex sentence may be true. According to this semantic picture, sentences (1.1)–(1.4) express false propositions, for they are simple sentences with an empty subject expression. Because sentence (1.5) is simply the negation of (1.2), and because (1.2) expresses a false proposition, (1.5) will express a true proposition. Because a fictional operator is a nontruth-functional operator, it can consistently be maintained that (1.6) and (1.7) are both true, even though the embedded sentences are false. It is important to note that this suggestion is consistent with the compositionality principle; the proposition expressed by any of the above sentences is a function of the semantic values, *or lack thereof,* of the constituents of that sentence.

This view can be wedded to a theory of singular propositions. It can be maintained that declarative sentences containing a nonmentioned occurrence of a proper name express singular propositions. When a declarative sentence contains an occurrence of an empty name—a name with no semantic value—it will express a special kind of singular proposition: a gappy proposition. A gappy proposition has a propositional structure like any other singular proposition, but one of the positions is left unfilled.[8]

---

[7] Cf. T.J. Smiley, "Sense without Denotation," *Analysis,* xx (1960): 125–35.

[8] Cf. David Kaplan, "Demonstratives," in Joseph Almog, John Perry, and Howard Wettstein, eds., *Themes from Kaplan* (New York: Oxford, 1989), pp. 481–564; David Braun, "Empty Names," *Noûs,* xxvii (1993): 449–69; Fred Adams and Robert Stecker, "Vacuous Singular Terms," *Mind and Language,* ix (1994): 387–401; Graham Oppy, "The Philosophical Insignificance of Gödel's Slingshot," *Mind,* cvi (1997): 121–41; and Salmon, "Nonexistence."

Let me illustrate, using a notation that is hopefully self-explanatory. According to the gappy proposition view, (1.3) and (1.4) both generate the following gappy proposition: <GAP, being-a-detective>. Likewise, (1.1) and (1.2) generate the following proposition: <GAP, exists>. Because these propositions (like the sentences that express them) are simple, all such propositions are false. Sentence (1.5)—the negation of (1.2)—will generate the complex gappy proposition: <NOT <GAP, exists>>. Because the proposition expressed by (1.2) is false, the proposition expressed by (1.5) is true. Sentences (1.6) and (1.7) will respectively generate the following complex gappy propositions: <Fiction$_{SH}$ <GAP, being-a-detective>> and <Fiction$_{CC}$ <GAP, exists>>. It can consistently be maintained that both propositions are true. This is all as we should expect.

This popular answer to the problem outlined above has the advantage of agreeing with common sense insofar as it allows (i) that sentences containing an occurrence of an empty name are meaningful, that is, that they express a proposition, and (ii) that the proposition expressed by such sentences can have a (classical) truth value. But the view has the disadvantage of disagreeing with common sense insofar as the theory entails that any two simple sentences, differing only with respect to the empty name contained therein, will express the same proposition. As a consequence, an adherent of this view appears committed to the thesis that substitution of one empty name for another in nonextensional contexts will always preserve the truth value of those sentences. But this does not seem right. Consider again sentences (1.6) and (1.7):

(1.6) According to the novel, *A Christmas Carol*, Ebenezer Scrooge exists.
(1.7) According to the Conan Doyle stories, Sherlock Holmes is a detective.

And compare them to the following:

(1.8) According to the novel, *A Christmas Carol*, Sherlock Holmes exists.
(1.9) According to the Conan Doyle stories, Ebenezer Scrooge is a detective.

Recall that, if we accept the direct reference theorist's solution to the original problem, the embedded simple sentence in (1.6) expresses the very same proposition as that expressed by the embedded simple sentence in (1.8). Likewise, the embedded sentence in (1.7) expresses the same proposition as that expressed by the embedded sentence in (1.9). As a consequence, (1.6) and (1.8) should express the very same proposition, and therefore have the very same truth value; so too should (1.7) and (1.9). But nothing could be further from the

truth. While (1.6) is true, (1.8) is false; (1.7) is true, yet (1.9) is false.[9] This, I maintain, is the main difficulty empty names present the direct reference theorist.

The problem, though, is not a problem for the direct reference theorist alone. It is a difficulty for the descriptivist also. Or more precisely: there is a similar problem for any descriptivist who embraces the plausible thesis that names are rigid designators.

II. DESCRIPTIVISM AND THE RIGIDITY THESIS

Kripke, in his groundbreaking monograph *Naming and Necessity*,[10] argued that proper names, as they occur in natural languages, are rigid designators. A term *t* is a rigid designator if and only if *t* designates the very same object in every possible world in which *t* designates anything at all. This definition of rigidity, however, is multiply ambiguous; it allows us to distinguish a number of different senses in which a designating term can be rigid. We can make them explicit as follows:

> A term *t* is *vacuously* rigid if *t* designates no object in any possible world; *t* is *nonvacuously* rigid if it designates some object *o* in some possible world, and in no world designates an object other than *o*. If *t* designates *o* in every possible world, whether or not *o* exists there, we call *t* *obstinately* rigid. If *t* designates *o* in all and only those worlds where *o* exists, we call *t* *persistently* rigid. If *t* designates *o* in only some of the worlds in which *o* exists, and in no others, we call *t* *tenaciously* rigid.

Which notion of rigidity does the semantic theorist have in mind when she defends the thesis that names are rigid designators? She is

---

[9] Fictional contexts, therefore, are (or appear to be) opaque contexts. It is a well-known fact that opaque contexts present a problem for direct reference theorists. But the literature on the subject has tended to focus on propositional attitude contexts rather than fictional contexts, for it appears that in such contexts, substitution of names—that is, either empty names or referring names—with the same semantic value will not always preserve the truth of what is said. It might be thought, therefore, that empty names per se, do not present a special problem for direct reference theories; perhaps the problem presented above is merely an instance of a more general problem. Nonetheless, there are two reasons for thinking sentences containing an occurrence of an empty name uttered in such contexts are of special interest: (i) if our attention is focused on empty names rather than referring names, we notice a similar substitutivity problem arises for the descriptivist, and (ii) various responses proffered by direct reference theorists to substitutivity problems in propositional attitude contexts—cf. especially Salmon, *Frege's Puzzle* (Cambridge: MIT, 1986); and Soames, "Direct Reference, Propositional Attitudes, and Semantic Content," *Philosophical Topics*, xv (1987): 47–87—are less plausible when we consider empty names in fictional contexts.

[10] "Naming and Necessity," was first published in Gilbert Harman and Donald Davidson, eds., *The Semantics of Natural Language* (Boston: Reidel, 1972); reprinted in monograph form (with preface added) (Cambridge: Harvard, 1980). All page references are to the later edition; hereafter NN.

not claiming that names are tenaciously rigid. Proper names are not tenaciously rigid,[11] but there are other designators that are good candidates for this title. Consider the definite description 'the father of Saul Kripke'. If we embrace the plausible assumption that Kripke could not have had a different father, it might be supposed that this description is tenaciously rigid. The description 'the father of Saul Kripke' designates the same individual, Myer Kripke, in every possible world in which it designates anything at all. It does not designate him in those worlds in which he does not exist, and (perhaps) it does not designate him in those worlds in which he is not a father.

The semantic theorist is thus forced to choose between the thesis that proper names are obstinately rigid designators and the thesis that proper names are persistently rigid designators. The overwhelming majority of philosophers who discuss the issue suggest that the right notion of rigidity is the first: proper names are obstinately rigid designators.[12] They reason as follows. Consider the sentence 'David Kaplan might not have existed'. Since Kaplan is not a necessary existent, this sentence must be true (that is, express a true proposition). According to the accepted semantics for such statements, this sentence is true only if there is a possible world in which 'David Kaplan does not exist' is true. But if proper names were only persistently rigid—if 'David Kaplan' designated nothing in those worlds in which he did not

[11] Or: most names are not tenaciously rigid. David Lewis has pointed out to me that there is a case to be made for the hypothesis that some proper names are tenaciously rigid. Perhaps 'Elizabeth II' is tenaciously rigid, if that name (as used by us in this world) does not designate Elizabeth in worlds where she exists but never takes the throne. Perhaps the new names taken by bishops when they become popes, or the name 'Great Leader' as used to designate Kim Jong-Il of North Korea, are also tenaciously rigid.

[12] Cf. Kaplan, "Demonstratives," and "Bob and Carol and Ted and Alice," in Jaakko Hintikka, Julius Moravcsik, and Patrick Suppes, eds., *Approaches to Natural Language* (Boston: Reidel, 1973), pp. 490–518; Colin McGinn, "Rigid Designation and Semantic Value," *Philosophical Quarterly*, XXXII (1982): 97–115; Esa Saarinen, "How to Frege a Russell-Kaplan," *Noûs*, XVI, 2 (May 1982): 257–276; Salmon, *Reference and Essence* (Princeton: University Press, 1981); and A.D. Smith, "Rigidity and Scope," *Mind*, XCIII (1984): 177–93, and "Semantical Considerations on Rigid Designation," *Mind*, XCVI (1987): 83–92. There are, however, two notable exceptions. Kripke categorically denies that the view of proper names that he intended to defend is the view that names are obstinately rigid. But he does not, as is often maintained, mean to suggest that they are persistently rigid either. Kripke intended to be deliberately neutral on this issue. Kripke's views on this matter are made clear in note 21 of NN, and in Kaplan, "Demonstratives," pp. 569–71. Robert Steinman—"Kripke Rigidity versus Kaplan Rigidity," *Mind*, XCIV (1985): 431–42—does defend the view that proper names are persistently rigid and not obstinately rigid. A good response to Steinman is to be found in Smith, "Semantical Considerations on Rigid Designation."

exist—then the sentence 'David Kaplan does not exist' would not be true at any world. Conclusion: proper names are obstinately rigid.[13]

It is worth noting, however, that there is a further way of characterizing the notion of rigidity to be found in the literature, a notion that is quite distinct from the three notions of rigidity defined above. It may be stated as follows:

> A term t is *insularly* rigid if and only if an utterance of a simple sentence S that contains a nonmentioned occurrence of t is such that t designates the same individual as it would if S appeared within the context of a modal operator.

It is important to note that insular rigidity is not equivalent to any of the three other notions of rigidity defined above. Nor is it a way of disambiguating Kripke's notion. Indeed, Kripke is explicit, in a number of places, that insular rigidity is a much weaker notion than the notion of rigidity that he had in mind. Nonetheless, many descriptivists have embraced this weaker notion, suggesting that it is consistent with the evidence for the rigidity hypothesis, and therefore any designator that meets this condition is a good enough candidate for the name 'rigid designator'.

Kripke mounts a concerted attack against the 'Frege-Russell' theory of names in *Naming and Necessity*. There he proffers a variety of arguments designed to persuade us of the untenability of such theories. Perhaps the most influential of these—the modal argument—can be reformulated as follows[14]:

Let 'N' be some name and 'the F' a description that uniquely picks out N—a purported candidate for the meaning of 'N'.

(i)   'The F might not have been an F' is ambiguous.
(ii) If 'N' were used to mean 'the F', then 'N might not have been an F' would also be ambiguous.
(iii) But 'N might not have been an F' is unambiguously true.
(iv) Therefore, 'N' is not used to mean 'the F'.[15]

---

[13] Of course, such an argument hardly seems compelling in light of the considerations outlined in the previous section. If 'Kaplan' designates nothing in worlds in which he does not exist, and as a consequence 'Kaplan exists' expresses a gappy proposition in those worlds, then the proposition would be false in those worlds, and moreover its negation would be true.

[14] It is interesting to note that Kripke himself does not advance an argument of this form. However, many commentators on *Naming and Necessity*—influenced by Brian Loar, "The Semantics of Singular Terms," *Philosophical Studies*, xxx (1976): 353–78—have casually attributed this argument to him.

[15] The argument requires some clarification. If we are to assume that the particulars that are named can possess essential properties, the uniquely identifying description associated with the name had better not pick out some such property. But as Kripke observes, this is hardly ever the case. It is rare (if not impossible) that we know enough

The objection can be applied to those who adopt a theory in which names are abbreviations of *a set or cluster* of descriptions.[16] According to the cluster theory, it might turn out that some of the descriptions associated with a name are false, and yet the name still refers. But for reference to occur it is necessary that some critical number of associated descriptions be true—or almost true. Kripke, quite rightly, has qualms about this modified descriptivism. Consider the following:

> Most of the things commonly attributed to Aristotle are things that Aristotle might not have done at all. In a situation in which he didn't do them, we would describe that as a situation in which *Aristotle* didn't do them.... Not only is it true *of* the man Aristotle that he might not have gone into pedagogy; it is also true that we use the term 'Aristotle' in such a way that, in thinking of a counterfactual situation in which Aristotle didn't go to *any* of the fields and do *any* of the achievements we commonly attribute to him, still we would say that was a situation in which *Aristotle* did not do these things (NN 61–62).

The modal argument against descriptivism is intimately associated with Kripke's doctrine that names are rigid designators. How so? It should be noticed that ordinary definite descriptions can give rise to scope ambiguities in modal contexts. That is why the first premise of the argument is true. Names, on the other hand, do not inherit the same kind of ambiguity. That is why the third premise of the argument is true. That names do not give rise to the same kind of ambiguity is entailed by the fact that names are rigid designators.

Some philosophers, in response to Kripke's argument, maintain that while the modal argument rationally compels the semantic theorist to embrace the thesis that names are rigid designators in some sense, it in no way counts against descriptivism. The semantic evidence is consistent, it is suggested, with the thesis that names are merely insularly rigid designators. If this diagnosis of the semantic evidence in support of the rigidity thesis is right, the descriptivist has at hand a response to the objection Kripke and others have mounted against their semantic theory.

The response runs as follows. Kripke is right to maintain that names are rigid designators; but they are only insularly rigid. This is because names are equivalent to descriptions that are semantically required to take wide scope over modal operators occurring in the same sentence. Thus, Kripke's objection does not get a hold. If there really is a

---

about the essential properties of any given individual that our related description will refer uniquely to that individual.

[16] Searle, "Proper Names," *Mind*, LXVII (1958): 166–73.

convention in natural language that proper names, unlike definite descriptions, always take wide scope over modal operators that occur in the same sentence, then the second premise of the modal argument is false. Call this view *wide-scope descriptivism*.[17]

It seems that this is an adequate response to the modal argument, as it is stated above. Nonetheless, it has been maintained by Kripke and others[18] that the argument can be modified in such a way as to make scope distinctions irrelevant. The modified version of the argument runs as follows:

Let '*N*' be some name and 'the *F*' a description that uniquely picks out *N*—a purported candidate for the meaning of '*N*'.

(i)   'The *F* is an *F*' is true in every world containing a unique *F*.
(ii) If '*N*' were used to mean 'the *F*', then '*N* is an *F*' would also be true in every world containing a unique *F*.
(iii) But '*N* is an *F*' is not true in every world containing a unique *F*.
(iv) Therefore, '*N*' is not used to mean 'the *F*'.

This version of the modal argument has a (supposed) advantage over the initial version of the argument insofar as the sentences we are asked to evaluate are simple sentences, and therefore cannot give rise to the scope ambiguities on which the initial version of the argument traded. Thus, the response to the initial version of the modal argument proffered by the wide-scope descriptivist cannot be applied to the modified version of the argument. The reason, it is suggested, is that names, unlike descriptions, are obstinately rigid, and not merely insularly rigid. Thus, descriptivism is false.

Whether or not this modified version of the modal argument decisively refutes wide-scope descriptivism is a matter of some controversy. Nonetheless, I do not care to adjudicate between the two sides here.[19]

---

[17] This response to Kripke has been proposed or embraced by an overwhelming number of philosophers sympathetic to the descriptivist position—see, for example, B. Brody, "Kripke on Proper Names," *Midwest Studies in Philosophy*, II (1977): 75–80; Michael Dummett, *Frege: Philosophy of Language* (Cambridge: Harvard, 1973); Brian Garrett, "More on Rigidity and Scope," *Logique et Analyse* (1984): 97–102; Loar; David Sosa, "Rigidity in the Scope of Russell's Theory," *Noûs*, XXXV (2001): 1–38; and Paul Yu, "The Modal Argument against Description Theories of Names," *Analysis*, XL (1980): 208–09. For criticism of this view, see NN 11–12; and Soames, "The Modal Argument: Wide Scope and Rigidified Descriptions," *Noûs*, XXXII (1998): 1–22.

[18] See the preface to NN, and James Hudson and Michael Tye, "Proper Names and Definite Descriptions with Widest Possible Scope," *Analysis*, XL (1980): 63–64.

[19] Some philosophers have responded to arguments like this by attempting to show that the premises of the argument are equivalent to sentences which do give rise to scope ambiguities, and so therefore the argument is unsound for exactly the same reason that the initial argument is (cf. Graeme Forbes, *Languages of Possibility* (New York: Blackwell, 1989); Sosa; and Yu).

Instead, I merely want to point out that even if this argument were successful against the wide-scope descriptivist, it would not decisively refute other varieties of descriptivism. Against Kripke, a descriptivist can still maintain that the relevant descriptions can be rigidified in the appropriate way so as to avoid any commitment to the conclusion of this new version of the modal argument. How might the descriptivist do this? Let me review one such suggestion.

The descriptivist I have in mind will suggest that the descriptions offered by Russell, Frege, and others be modified in such a way as to index them to a possible world. Perhaps the most straightforward way to do this is to incorporate an actuality operator into the relevant descriptions. Thus, it might be said that 'Aristotle', for example, means something like the *actual* pupil of Plato and the *actual* teacher of Alexander the Great. This new kind of description is a kind of obstinately rigid designator. So where does Kripke's modal argument go wrong? If the definite description used in either version of the argument is a world-indexed description, premise (i) will be false. Call this view *world-indexed descriptivism*.[20]

We are now in a position to demonstrate that if the descriptivist seeks to accommodate Kripke's modal intuitions—and thereby respect the evidence for the rigidity of proper names—she will, somewhere along the line, find herself in trouble. Where we locate the trouble will depend on which variety of descriptivism—world-indexed descriptivism or wide-scope descriptivism—is under consideration.

III. WORLD-INDEXED DESCRIPTIVISM

Consider the following three theses about Santa Claus:

(3.1) Santa Claus does not exist.
(3.2) Santa Claus might have existed.
(3.3) It is possible that Santa Claus has a beard, lives at the north pole, climbs down chimneys to give children presents on Christmas morning, and so forth.

As is well known, the descriptivist (like the direct reference theorist) can account for our transparent intuition that (3.1) is true. She has

---

[20] This response to Kripke is embraced by philosophers such as Leonard Linsky, *Names and Descriptions* (Chicago: University Press, 1977); Plantinga; and Frank Jackson, "Reference and Description Revisited," *Philosophical Perspectives*, XII (1998): 201–18. A thorough and sympathetic presentation of the idea is given in Jason Stanley, "Names and Rigid Designation," in Bob Hale and Crispin Wright, ed., *A Companion to the Philosophy of Language* (Cambridge: Blackwell, 1997), pp. 555–85. For criticism of this view, see Salmon, *Reference and Essence*, pp. 26ff. and Soames, "The Modal Argument."

at hand something like the following plausible paraphrase of (3.1) where 'F!' is a predicate like 'is a *unique* big fat jolly man who has a beard and lives at the north pole and so forth'; 'A' is the modal operator 'actually'[21]:

> (3.1A) $\sim\exists x\,(AF!x)$

Moreover, it is clear that the descriptivist can paraphrase a vast array of other sentences containing empty names in an intuitively plausible way. This advantage, though, has all too often been overemphasized. Let us see why.

Consider first world-indexed descriptivism. The adherent of this view will paraphrase (3.2) as something like,

> (3.2A) $\Diamond\exists x\,(AF!x)$

On this account, (3.2) is false. If it is false that *in the actual world*, there is an x that has a beard, lives at the north pole,..., then it is also false that there is a possible world such that: *in the actual world* there is an x that has a beard, lives at the north pole,.... Santa Claus, then, could not have existed according to world-indexed descriptivism.

It seems to me that this is a prima facie cause for concern. It is a commonplace that Santa Claus might have existed.[22] The modality here is not an epistemic one. We continue to be disposed to assert (3.2) once we learn that Santa does not in fact exist. There may of course be some philosophers, deeply entrenched in their own theoretical views, who have forgotten that this is the common-sense view. But it should be clear from Kripke's own comments on the subject that to deny (3.2) does some violence to our intuitions. Consider the following:

---

[21] The open formula $F!x$ is simply shorthand for $(Fx\,\&\,\forall y\,(Fy \rightarrow x = y))$ and can be read as 'x is a unique F'. See J. Crossley and I.L. Humberstone, "The Logic of 'Actually'," *Reports on Mathematical Logic*, VIII (1977): 11–29, for a detailed explanation of the logical behavior of the actuality operator: 'A'. Of course the wide-scope descriptivist's translation of (3.1) will omit all occurrences of 'A' (cf. paraphrase (3.1B)).

[22] In response to the argument I present below against world-indexed descriptivism, it has sometimes been suggested to me that (3.2) should be taken to mean something like:

(3.2*) 'Santa Claus' might have referred.

If it is, the trouble disappears. But the suggestion seems implausible to me, for on this view, negative existentials that "use" different names cannot have the same meaning—they automatically express different propositions. Intuitively, though, it seems as though the sentence 'Lè Pere Noël n'existe pas' can be used to express exactly the same proposition as that expressed by an utterance of (3.1).

Holmes does not exist, but in other states of affairs, he would have existed.[23]

And later, after a change of opinion:

Some of the views I have are views which may at first glance strike some as *obviously wrong*. My favorite example is this (which I probably won't defend in these lectures—for one thing *it doesn't ever convince anyone*): It is a common claim in contemporary philosophy that there are certain predicates (and names) which, though they are in fact empty—have null extension—have it as a matter of contingent fact.... [A]n example often given is the example of a unicorn. So it is said that though we have all found that there are no unicorns, of course there might have been unicorns. Under certain circumstances there would have been unicorns. And this is an example of something I think is not the case... Now I don't know if I'm going to have time to defend this particular view, *but it's an example of a surprising one....* So, *some of my opinions are somewhat surprising* (NN 23–24, my emphasis).

But commonplaces can be challenged. We often find good reasons to reject common-sense intuitions, reasons that are independent of our current theoretical concerns. It is ironic that Kripke himself has provided the adherent of world-indexed descriptivism with what might seem to be the required justification.[24]

Kripke puts forward a number of important considerations that are meant to persuade us that sentences like (3.2) are in fact false. Kripke's discussion focuses for the most part on sentences that contain a nonreferring natural kind term (like 'unicorn'). But the most salient of these considerations pertain to sentences that contain an instance of a nonreferring proper name. Kripke suggests that no matter what cluster of properties we have in mind when we use terms like 'unicorn' or 'Santa Claus', if the term does not refer to any actual object it cannot refer to any possible object either. Why? Because it is much harder—if not impossible—to find a description that uniquely picks out a merely possible individual (or natural kind) than it is to find one that picks out an actual individual (or natural kind). Any number of nonactual characters could fit the description found in the Santa Claus story. Indeed, there may be a single world in which a plurality of individuals are such that they live at the north pole, have a beard, own flying reindeer, give out presents early on Christmas morning,

---

[23] Kripke, "Semantical Considerations on Modal Logic," *Acta Philosophica Fennica*, XVI (1963): 83–94, here p. 85.
[24] Kaplan, in "Bob and Carroll and Ted and Alice," makes essentially the same point in Appendix XI, pp. 505–08.

and so forth. But recall that the descriptivist and direct reference theorist alike are in agreement that names are rigid designators. Thus names can refer to only one (possible) individual. Because there is no fact of the matter to appeal to when deciding which of these hypothetical characters would have been Santa Claus, it seems wrong to suggest that Santa might have existed at all.

Kripke's argument might seem compelling, and the adherent of world-indexed descriptivism may well be happy to take it on board. No matter what one thinks of the argument, though, it is not meant to show that it is obvious that (3.2) is false. Indeed, Kripke himself says that his "opinions are somewhat surprising." Thus, there remains a residual worry for the descriptivist, a worry that does not pertain to the direct reference theorist.

One of the main tenets of descriptivism, it will be recalled, is that names have senses and whenever we are in a position to use a name we must have the sense associated with it in mind. As has already been noted, the sense of a name like 'Santa Claus', according to world-indexed descriptivism, can be captured by something like the following description: 'The *actual* jolly fat man who lives at the north pole, has a beard, and so forth'. The worry we should have, then, is that once we come to realize that there is no individual who fits that description, and therefore that Santa Claus does not in fact exist, we should be in a position to infer immediately that Santa could not possibly have existed. To fail to make this observation—if we really had this description in mind whenever we used the name 'Santa Claus'—would be to suffer from some rational failing or other, perhaps because we do not understand the seemingly obvious logical relation that stands between sentences (3.1) and (3.2), perhaps because we do not pay attention to (or understand) our own use of the modal operator 'actually'. But most people are disposed to assert both (3.1) and (3.2), and it seems to me that we should be suspicious of any theory that attributes such widespread stupidity.[25]

[25] An anonymous referee has pointed out that this objection to the wide-scope descriptivist might not be as costly as I make it appear. It is commonly accepted, for example, that no human is logically omniscient; that is, humans do not know all of the logical consequences of what they know. But this is not a reason for attributing widespread stupidity; to be ignorant of the logical entailments of one's beliefs is not a way to be dim-witted or irrational. It might be thought, therefore, that the wide-scope descriptivist can take succor from this observation. Perhaps our failure to notice the necessity of Santa Claus's nonexistence after recognizing that there is no Santa is simply an instance of this general phenomenon. It seems to me, however, that my objection is relevantly different in two respects. First, if the world-indexed descriptivist is right about the content of our beliefs, most people who know that Santa does not exist not only fail to believe a logical consequence of that belief, they also believe the negation of such a consequence. Second, if the wide-scope descripti-

IV. WIDE-SCOPE DESCRIPTIVISM

It might therefore be hoped that trouble can be avoided if we adopt instead a kind of wide-scope descriptivism. An adherent of this kind of descriptivism will be in agreement with common sense insofar as she is committed to the truth of theses (3.1) and (3.2). She will presumably paraphrase each as follows:

(3.1B) $\sim\exists x\ (F!x)$
(3.2B) $\diamondsuit\exists x\ (F!x)$ [26]

There is, however, a problem for an adherent of this variety of descriptivism, a problem analogous to that mounted against the world-indexed descriptivist.

If it is possible that Santa Claus exists, then it should also be possible that Santa Claus is *an F*, where '*the F*' is the description capturing the meaning of the name 'Santa Claus'. If the second thesis about Santa is true, the third should be true also. For it is analytic that it is possible that there exists an *F* only if it is possible that there exists an *F* that is *F*.

The wide-scope descriptivist, though, cannot agree. Recall that the wide-scope descriptivist seeks to endorse the rigidity thesis by maintaining that the logical form of sentences containing a proper name and a modal operator are such that the description (of which the name is an abbreviation) invariably takes wide scope. Thus, her translation of (3.3) should be as follows:

(3.3B) $\exists x\ (F!x\ \&\ \diamondsuit Fx)$

But (3.3B) is clearly false. We have learned that there is no individual that is picked out by the Santa Claus description, yet (3.3B) says there is. The account must therefore be rejected.

The arguments outlined in the last two sections spell trouble for the descriptivist. But the arguments are not knock-down arguments, and the keen reader will no doubt anticipate a number of potential escape routes available to the descriptivist. In the sequel I will consider two such rejoinders, one inspired by Russell, the other by Frege. My

---

vist were right, the consequence should be a trivial one. And it is a minimal requirement of rational agents that they recognize the trivial consequences of their beliefs, even if they are not required to recognize every logical consequence.

[26] Notice that there is no room for any scope ambiguities here, for 'existence' is not treated as a predicate. If, however, it were, sentence (3.2) would instead be paraphrased as follows: $\exists x\ (F!x\ \&\ \diamondsuit Ex)$. Because there is in fact no *F*, an adherent of this view is thereby committed to the view that (3.2) is false. Thus, she is vulnerable to exactly the same objection mounted against the world-indexed descriptivist in section III.

task here is not evaluative; my aim is merely to point out that however the descriptivist might respond, a parallel and equally plausible response is open to the direct reference theorist.

V. THE HYBRID VIEW

If what I have said thus far is correct, there is trouble in the offing for the descriptivist who is committed to the thesis that all names—both those that refer and those that do not—are rigid designators. But perhaps she need not commit herself to such a thesis; perhaps all that the preceding discussion shows is that empty names cannot be rigid designators even if ordinary referring names are.

The adherent of a hybrid variety of world-indexed descriptivism will suggest that a referring name will be captured by a rigidified description of the form 'the *actual F*', whereas a nonreferring name will be captured by a nonrigidified description of the form 'the *F*'. Those who embrace a hybrid variety of wide-scope descriptivism, on the other hand, will suggest that the logical form of any sentence containing a proper name and a modal operator will vary depending on whether or not the name refers. If it does, the description associated with the name will take wide scope; if it does not, the description will take narrow scope.

The semantic theorist who accepts this kind of hypothesis is in a position to reconcile her theory with some of our firmest intuitions about proper names. She will agree with all of Kripke's modal intuitions, except when it comes to the case of empty names. When we restrict our attention to the case of empty names only, premise (3) of Kripke's modal argument—the premise that any sentence of the form '*N* might not have been the *F*' is unambiguously true—loses much of its plausibility. For when we come to consider ordinary and mundane varieties of empty names it is not unreasonable to suppose that all (or most) of the descriptions we associate with these names are, so to speak, essential to the characters in question.

Moreover, the proponent of this style of account will be able to accommodate our intuitions that Santa Claus does not exist, that it is possible that he exists, and that if he had existed, he might have been a kind fat jolly man with a beard. An adherent of either variety of descriptivism about referring names—wide-scope descriptivism or world-indexed descriptivism—who also advocates the hybrid view about proper names in general, will suggest that theses (3.1)–(3.3) have the following logical form:

(3.1B) $\sim \exists x \, (F!x)$
(3.2B) $\Diamond \exists x \, (F!x)$
(3.3H) $\Diamond \exists x \, (F!x \,\&\, Fx)$

On this account, then, theses (3.1)–(3.3) are all true. Indeed, thesis (3.3) is analytic. This is all as we should expect.

Such a proposal commits us to the view that the difference between the semantic value of an empty name and that of a referring name is a difference in kind. This difference is postulated because of the problems outlined in section III and section IV. Having made such a concession to the descriptivist, however, we must be prepared to make a similar allowance for the direct reference theorist.

A number of philosophers sympathetic to the theory of direct reference, and bothered by the difficulties empty names present this theory, have suggested that so-called empty names are not really names at all.[27] Instead, such terms are treated as disguised descriptions. Because they are not singular terms, it is maintained, the Millian is not responsible for interpreting sentences containing an occurrence of an empty 'name'.[28]

Whether we choose to call terms like 'Santa Claus', 'Vulcan', and 'Sherlock Holmes' names or not seems of little relevance. The important point to note in the present context is that, on such a view, the semantic value of these terms is of a significantly different kind to that of referring names. Thus, Currie and Ryckman both offer the direct reference variant of the hybrid view. To illustrate, consider the following matrix, representing the alternative views.

	Referring Names	Empty Names
Descriptivism	Rigidified Description	Flaccid Description
Direct Reference	An Individual	Flaccid Description

Whatever the merits and disadvantages of these views, their respective treatment of empty names is exactly the same in both cases; both stand or fall together on this issue.[29]

---

[27] See, for example, Ryckman; and Gregory Currie, *The Nature of Fiction* (New York: Cambridge, 1990), chapter 4.

[28] Russell's official view about proper names was very close to the theory propounded by the hybrid variety of direct reference theory. For Russell, though, the dichotomy is slightly different; instead of referring names, Russell considers the class of names we are absolutely certain refer. These names, suggests Russell, are directly referential. All others are abbreviated definite descriptions.

[29] Lewis has suggested to me that there is an alternative kind of hybrid view that might be proposed by the descriptivist which would evade the problems outlined in section III and section IV. Lewis suggests that descriptions generally, and proper names in particular, are once and for all ambiguous between a rigidified and unrigidified sense. We therefore have to rely on facts about the conversational context to determine which kind of proposition the speaker intends to express. This kind of view is significantly different than the position I outline above insofar as the adherent of this kind of view has an easy response to Kripke's modal argument (outlined in section II). Unlike the descriptivist of any variety considered thus far, she will deny premise (3) of the modal argument—the premise that sentences like 'Aristotle might

VI. THE REALIST VIEW

Recall that a descriptivist says that names are just abbreviated definite descriptions and so a sentence containing a proper name has the same logical structure as a sentence containing a definite description. We may, however, distinguish two main approaches to the problem of how best to analyze the logical structure of sentences containing a definite description.

The first, Fregean, approach treats definite descriptions as singular terms, so that '*a* is *G*' and 'the *F* is *G*' genuinely have the same sentential structure, though the latter involves a complex singular term. One virtue of this account is that the logical form of a sentence like 'the cat in the hat came to play' reflects its surface grammatical form.

The second, Russellian, approach denies that definite descriptions are singular terms. Instead, definite descriptions are treated as a species of singular restricted quantifier, binding variables rather than substituting for them, so that '*a* is *G*' and 'the *F* is *G*' differ in logical form. More precisely, a definite description of the form 'the *F* is *G*' is paraphrased as follows:

(6.1) $\exists x \, (F!x \, \& \, Gx)$

When outlining the difficulties empty names present the descriptivist, I have assumed (with Russell) that descriptions are a species of singular restricted quantifier rather than singular terms. This is not because I believe the objections do not have force against the Fregean view. The reason for preferring Russell's approach over Frege's starts from the observation that because the Fregean approach treats definite descriptions as singular terms, it seems to face a similar problem to that outlined in section I against the direct reference theorist. Like the direct reference theorist, Frege is not committed to the view that sentences containing an occurrence of an (apparently) empty singular term have no meaning. But Frege, at first blush, seems to be committed to the view that all such sentences have no (classical) truth value. For according to Frege, a declarative sentence has a (classical) truth value only if all singular terms contained therein refer.

---

not have been a philosopher' are unambiguously true. Moreover, a descriptivist of this variety will not have any difficulty accounting for our intuitions about sentences containing an occurrence of an empty name. It is important to note, however, that if one finds this kind of view plausible, a similar response to the problem of empty names is open to the direct reference theorist. She can likewise suggest that proper names are once and for all ambiguous (between a descriptivist and directly referential sense).

Frege maintained that the problem could be solved in a variety of different ways. But in order to preserve our intuition that all such sentences express a proposition with a classical truth value, Frege suggested that a singular term "shall in fact designate an object, and that no new sign shall be introduced as a proper name without being secured a reference" (*op. cit.*, p. 70). The way Frege proposes to do this is to insist that an arbitrary object be provided if necessary. He says that a singular term "must always be assured of reference by means of a special stipulation, for example by the convention that 0 shall count as its reference when the concept applies to no object or more than one" (*op. cit.*, p. 70).

Whatever the merits of this arbitrarily-chosen-object view, it is inadequate as an account of the semantics of (so-called) empty names in natural language.[30] There are at least two problems with the approach: (i) the view appears to give the wrong distribution of truth values to the following kinds of sentences: 'Sherlock Holmes is a number', 'The name "Bart Simpson" has a referent', 'Johnny believes that Santa exists because he believes that the number 0 exists', all of which are (or appear to be) true according to this hypothesis, and, perhaps more importantly, (ii) an adherent of this approach appears committed to the view that substituting one "empty" term for another in any sentence *S* can never change the truth value of the proposition expressed by *S*, for the referent in each case will be the same.[31]

One way out of this difficulty is to suggest first that there is an abundance of such objects and second, that the individuals to which these so-called empty names refer is no arbitrary matter. Perhaps such names pick out various Meinongian nonexistent objects, or perhaps they denote abstract particulars, or perhaps they refer to some other

[30] Frege, of course, never maintained that this was a way of explaining our intuitions about empty names as they occur in natural languages. It was a technical solution to a technical difficulty, that is, a suggestion for side-stepping this problem when we come to construct an artificial language. Nonetheless, other philosophers, influenced by the work of Frege (for example, Rudolf Carnap, Alonzo Church, Michael Dummett, and W.V. Quine), have since arrogated to themselves the proposed solution, and put it forward as a way of understanding natural languages.

[31] Frege himself solved the problem, at least in part, by suggesting that in some nonextensional contexts names and descriptions change their mode of reference. Instead of referring to an individual, such terms come to refer to their sense. I do not consider the merits of such a proposal here, however, because (i) it does not explain why sentences like 'Sherlock Holmes is a number' have the truth value they do; and (ii) this proposal, by itself, is of no help dodging the problems outlined in section III and section IV. For an interesting discussion about how such a suggestion might help the direct reference theorist, though, see Oppy.

kind of thing altogether.[32] But no matter how the realist's thesis is spelled out, we have a solution to the problem at hand. Santa Claus might have existed (because there is such an entity); and if he had, he would have been a kind fat jolly man (or at least, he would have been so represented). Moreover, we can leave room for a sense in which claims like 'There is no Santa Claus' are true, so long as we understand such quantification as implicitly restricted.

Embracing this kind of realism about these supposedly nonexistent entities also brings into line the Fregean theory with the natural view that sometimes substituting one nonreferring term for another in a given sentence can change the truth value of that sentence. But recall that this was the very same problem facing the direct reference theorist. Consider again sentences (1.7) and (1.9):

(1.7) According to the Conan Doyle stories, Sherlock Holmes is a detective.
(1.9) According to the Conan Doyle stories, Ebenezer Scrooge is a detective.

If the names 'Ebenezer Scrooge' and 'Sherlock Holmes' are genuinely nonreferring, then they have the same semantic value—none. And if two names have the same semantic value, substituting one name for the other in any given sentence could not change the truth value of the sentence. As a consequence, (1.7) and (1.9) should express the very same proposition, and therefore have the very same truth value. But this is clearly not the case. While (1.7) is true, (1.9) is clearly false.

One way out of the difficulty is to deny the assumption that gives rise to the problem, the assumption that 'Sherlock Holmes' and 'Ebenezer Scrooge' have the same semantic value. The realist does just that by suggesting that such names are not denotationless terms, but names that refer to two very different abstract or Meinongian individuals.

---

[32] For independent reasons to adopt this option, see Robert Howell, "Fictional Objects: How They Are and How They Aren't," *Poetics*, VIII (1979): 129–77; Kripke, "Reference and Existence," Unpublished Manuscript: Presented as The John Locke Lectures (1973); Peter Lamarque, "Fiction and Reality," in Lamarque, ed., *Philosophy and Fiction: Essays in Literary Aesthetics* (Aberdeen: University Press, 1983), pp. 52–72; Jerrold Levinson, "Making Believe," *Dialogue*, XXXII (1993): 359–74; Terence Parsons, *Nonexistent Objects* (New Haven: Yale, 1980); Peter van Inwagen, "Creatures of Fiction," *American Philosophical Quarterly*, XXIV (1977): 299–08; and Nicholas Wolterstorff, "Characters and their Names," *Poetics*, VIII (1979): 101–27. Such authors tend to focus on the case of *fictional* names. But it seems to me that the considerations used by such authors to support the thesis that such names refer apply equally well, for example, to the case of names that were introduced by mistake, when someone honestly thought they referred (for example, the name 'Vulcan' which was introduced by Le Verrier to designate the unobserved planet which caused the deviations in Mercury's orbit predicted by Newton's laws of motion and gravity).

VII. CONCLUSION

And so I rest my case, the case in favor of the hypothesis that empty names do not present a special problem for the direct reference theorist; to the extent that the direct reference theorist has a problem with empty names, the descriptivist has one also. The arguments presented, though, are not decisive. Perhaps there is a natural way to dodge the difficulty available only to the descriptivist. The fact that I have not noticed it may reflect nothing more than a lack of ingenuity on my part. But, at the very least, it should be clear that this all too common presupposition can no longer be taken for granted.

STUART BROCK

Victoria University of Wellington

DIALECTICAL DELICACIES IN THE DEBATE ABOUT FREEDOM AND ALTERNATIVE POSSIBILITIES*

It is widely assumed that the freedom necessary for moral responsibility requires alternative possibilities.[1] This intuitively compelling idea, that responsibility presupposes that we have open pathways branching into the future, is encapsulated in the *principle of alternative possibilities* (PAP):

> PAP: A person is morally responsible for what she has done only if she could have done otherwise.[2]

There are, however, various ways to challenge PAP.[3] One invokes famed "Frankfurt-type" examples, so named after Harry Frankfurt's presentation of them. Such examples contain a fail-safe mechanism that does not play any role in the relevant agent's actual deliberations, choices, and behavior but that ensures that the agent deliberates, chooses, and behaves in just the way in which he does. In his original examples, Frankfurt relied upon a counterfactual intervener who monitored the agent and intervened in the agent's behavior should the agent show any tendency to do other than what the intervener desires the agent to do.[4] In subsequent refinements, if intervention is called for, the intervener relies on a "prior-sign" as a cue for such intervention. To illustrate, consider this *prior-sign Frankfurt example*:

---

[1] Philosophers as disparate as Hume and Kant shared this view. For just a few of the many contemporary examples, see Carl Ginet, *On Action* (New York: Cambridge, 1990); Robert Kane, *The Significance of Free Will* (New York: Oxford, 1996); David Lewis, "Are We Free to Break the Laws?" *Theoria*, XLVII (1981): 113–21; and Peter van Inwagen, *An Essay on Free Will* (New York: Oxford, 1983). As we shall make clear, some argue that moral responsibility requires freedom, but not freedom involving alternative possibilities. A minority position challenges the very basic assumption that moral responsibility requires freedom. For example, in *Moral Responsibility and Persons* (Philadelphia: Temple, 1992), Eugene Schlossberger rejects the view that responsibility is a function of the agent's acting freely—see, for example, pp. 4–7, 117–18. Similar ideas are suggested in Nomy Arpaly's *Unprincipled Virtue* (New York: Oxford, 2003), pp. 149–79.

[2] See Harry G. Frankfurt, "Alternate Possibilities and Moral Responsibility," this JOURNAL, LXVI, 23 (December 4, 1969): 829–39, on p. 829.

[3] We shall focus upon Frankfurt's well-known approach. For an alternative, see, for instance, Michael J. Zimmerman, "Taking Luck Seriously," this JOURNAL, XCIX, 11 (November 2002): 553–76. Zimmerman argues that considerations of luck impugn PAP.

[4] See Frankfurt, pp. 835–36.

Mork believes that in his circumstances it is wrong for him to lie to prevent embarrassment. In spite of this, he decides to lie and lies anyway. As he has no excuse or justification for lying, it seems he is deserving of blame for lying. Unbeknownst to Mork, though, he could not have refrained from lying due to the presence of Mindy who has the power to read and control Mork's mind. Mindy wields this power partly in virtue of possessing the following knowledge. Had Mork been about to refrain from lying, he would have displayed some involuntary sign—a neurological pattern, N^*, in his brain—whereas if he had been about to lie, he would have displayed a different neurological pattern, N. Had Mindy detected N^*, she would have interceded in Mork's deliberations via direct stimulation of Mork's brain and would, in this way, have caused Mork to lie. But Mindy detects N, the reliable sign for Mindy that she need not show her hand at all. As Mork lies on his own, in the absence of Mindy's intervention, it seems highly reasonable that Mork acts freely and is morally blameworthy for lying despite not having alternative possibilities.

In light of such examples, Frankfurt and others have argued that alternative possibilities are *not* required for blameworthiness or responsibility in general. Call those who defend Frankfurt's argument *Frankfurt Defenders.*

A great deal turns upon Frankfurt's thought experiment.[5] If it is sound, one compelling argument for incompatibilism is defeated. This incompatibilist argument, the *Classical Incompatibilist Argument*, maintains that if determinism is true, no one can do otherwise; free will and moral responsibility require the ability to do otherwise; therefore, if determinism is true, no one has free will, and no one is morally responsible. Traditionally, compatibilists had wrestled with the Classical Incompatibilist Argument by challenging the premise that determinism rules out the ability to do otherwise. Frankfurt's approach offered compatibilists a way of conceding this premise while undermining this incompatibilist argument on other grounds. If, as Frankfurt argues, PAP is false, free will and moral responsibility do not require the ability to do otherwise.

Prior-sign examples have played such a crucial role in this debate because it has seemed to many that without prior-signs, Frankfurt's argument is transparently unconvincing. To illustrate, consider a

---

[5] Some terminology will prove helpful. Determinism is the thesis that, holding fixed the nonrelational facts of the past and the laws of nature, there is at any instant exactly one physically possible universe. Compatibilists believe that determinism is compatible with free will and moral responsibility. Incompatibilists hold that free will and moral responsibility are incompatible with determinism. Libertarians are incompatibilists who believe that determinism is false and that some human beings sometimes act of their own free will.

Frankfurt example that does not involve a prior-sign. Suppose that in the above example, Mindy is only able to intervene in Mork's conduct *after* Mork has decided (chosen, formed the intention) not to lie. Only in this way can Mindy ensure that Mork will lie (or, at least, utter a falsehood). If this is how the example goes, the advocate of PAP can simply point to the fact that when Mork freely lies on his own, even in Mindy's presence, he did have available a significant alternative possibility[6]; he could have decided (chosen, formed an intention) not to lie. So, the defender of PAP can argue, Frankfurt examples that do not involve prior-signs are not counterexamples to PAP. An effective counterexample must rule out any morally significant alternative within the voluntary control of the agent. Prior-sign examples, ex hypothesi, make it so that the ensuring conditions (that guarantee only one course of action) rule out alternatives prior to the initial moments of freely willed actions.[7]

I. THE DILEMMA DEFENSE AND A NEGLECTED HORN

One compelling objection to prior-sign examples is couched in terms of a dilemma.[8] If sign N is reliable, in the sense of being infallible, it can only be because states of the agent prior to the occurrence of the putatively free action are causally sufficient for this action (and the sign indicates this). But if that is the case, then a deterministic relation obtains between the prior-sign and subsequent behavior, and this begs the question against incompatibilists who believe that determinism is incompatible with freedom or responsibility. If, on the other hand, sign N is not infallible and is only reliable in some weaker sense, then an agent who acts freely in a Frankfurt example retains the ability to do otherwise when she acts on her own. The presence (or absence) of the prior-sign is consistent with the agent acting in a manner other than the manner in which she does. Call this defense of PAP the *Dilemma Defense*, and call those who advance it, *Dilemma Defenders*.[9]

---

[6] By a "morally significant alternative possibility" we mean to restrict our attention only to what some have called "robust" alternatives. Robust alternatives are alternatives that are within the scope of the agent's control and that can be used as a (partial) basis for explaining why an agent *is* morally responsible for her action (or in their absence, why an agent is not morally responsible for her action).

[7] This moment might be called the *locus* of a freely willed action, the baptismal point of human freedom or control.

[8] See Carl Ginet, "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing," *Philosophical Perspectives*, x (1996): 403–17; Kane, pp. 142–44; and David Widerker, "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities," *Philosophical Review*, CIV (1995): 247–61.

[9] Note that the Dilemma Defense seems only to be available to incompatibilists,

The Dilemma Defense has instigated an impressive literature. Most Frankfurt Defenders have attempted to respond only to the second horn of the dilemma. They have attempted to produce examples that both rule out all morally significant alternative possibilities and that do not assume that the agents in the examples are determined.[10] With one notable exception,[11] Frankfurt Defenders have ignored the first horn of the dilemma.

In this paper, we focus principally on this horn—the horn that objects that the prior-sign cannot presuppose a deterministic relation. Our primary aim is to bring into relief at least two separate dialectical contexts that naturally arise in thinking about its presumed significance. These contexts can be sorted out by fixing upon two hitherto unnoticed points. One concerns distinct interpretations of the first horn. Another concerns the different audiences at which Frankfurt's argument can be appropriately directed. We propose that one interpretation of the first horn, the *broad interpretation*, is most reasonably suited for a dialectic in which Frankfurt's argument is directed at a philosophically *committed audience*. We also propose that a very different interpretation of the first horn, the *narrow interpretation*, is most reasonably suited for a dialectic in which Frankfurt's argument is directed at an *undecided audience*.

Distinguishing between these dialectical dynamics is significant be-

---

since it is only they who are entitled to the objection that, given one horn of the proposed dilemma, Frankfurt examples might infelicitously presume a deterministic relation.

[10] See Ishtiyaque Haji, *Moral Appraisability* (New York: Oxford, 1998), pp. 30–41; David Hunt, "Moral Responsibility and Unavoidability," *Philosophical Studies*, XCVII (2000): 195–227, and "Freedom, Foreknowledge, and Frankfurt," in Widerker and Michael McKenna, eds., *Moral Responsibility and Alternative Possibilities* (Aldershot, UK: Ashgate, 2003), pp. 159–83; Alfred Mele and David Robb, "Rescuing Frankfurt-Style Cases," *Philosophical Review*, CVII (1998): 97–11, and also "BBs, Magnets, and Seesaws: The Metaphysics of Frankfurt-Style Cases," in Widerker and McKenna, eds., pp. 127–37; McKenna, "Robustness, Control, and the Demand for Morally Significant Alternatives," in Widerker and McKenna, eds., pp. 201–17; Derk Pereboom, "Alternative Possibilities and Causal Histories," *Philosophical Perspectives*, XIV (2000): 119–37, *Living without Free Will* (New York: Cambridge, 2001), and "Source Incompatibilism and Alternative Possibilities," in Widerker and McKenna, eds., pp. 185–99; and Eleonore Stump, "Libertarian Freedom and the Principle of Alternative Possibilities," in Daniel Howard-Snyder and Jeff Jordan, eds., *Faith, Freedom, and Rationality: Essays in the Philosophy of Religion* (Lanham, MD: Rowman and Littlefield, 1996), pp. 73–88, and "Moral Responsibility without Alternative Possibilities," in Widerker and McKenna, eds., pp. 139–58.

[11] See John Martin Fischer, "Recent Work on Moral Responsibility," *Ethics*, CX (1999): 93–139, "Frankfurt-Type Examples and Semi-Compatibilism," in Kane, ed., *The Oxford Handbook of Free Will* (New York: Oxford, 2002), pp. 281–308, and "Free Will and Moral Responsibility," in David Copp, ed., *Oxford Handbook of Ethical Theory* (New York: Oxford, forthcoming).

cause a response to the first horn depends crucially upon which dynamic is presupposed. A response seemingly effective on one interpretation misses the mark on the other. Furthermore, uncovering these different dialectical contexts will help to unearth and assess what Dilemma Defenders are centrally committed to. We propose that, on one interpretation, the Dilemma Defender is obliged to entertain the possibility that determinism rules out free will for reasons that have nothing to do with alternative possibilities. We propose that, on a second interpretation, it is *not* evident why the Dilemma Defender should assume that she is entitled to claim that the first horn is question begging.

II. TWO INTERPRETATIONS, TWO AUDIENCES, AND TWO DIALECTICAL SITUATIONS

The narrow interpretation of the Dilemma Defender's first horn is that, in those cases where the prior-sign assumes a deterministic relation, Frankfurt examples beg the question against the incompatibilist *because the deterministic relation expunges alternative possibilities.* It does so regardless of the other factors figuring in a Frankfurt example (factors such as counterfactual interveners). In contrast, on the broad interpretation, in those cases where the prior-sign assumes a deterministic relation, Frankfurt examples beg the question against the incompatibilist *because a deterministic relation obtains—simpliciter.* Understood this way, the concern is that, *independently of the importance of alternative possibilities,* Frankfurt-type examples that assume a deterministic relation cannot be used to convince *incompatibilists* that free will and moral responsibility do not require alternative possibilities. According to the incompatibilist, a *determined* agent is not free or morally responsible in the first place.

We now elaborate these interpretations. Beginning with the broad interpretation, suppose that, dialectically, the situation is this: the Frankfurt Defender means to persuade incompatibilists, who are already theoretically *committed* to a position regarding the free will problem.[12] In this dialectical situation, all that the Frankfurt Defender wants the *incompatibilist* to accept is that, if determinism *does* rule out the freedom required for moral responsibility, it is *not* by virtue of

---

[12] We ignore here the yet *further* dialectical context in which the Frankfurt defender means to persuade the committed compatibilist. This context has its own unique problems, but it does not center on contesting the assumption that an agent in a Frankfurt example is determined. For the best compatibilist resistance to Frankfurt examples, see Bernard Berofsky's "Classical Compatibilism: Not Dead Yet," in Widerker and McKenna, eds., pp. 107–26.

ruling out alternative possibilities. So the Frankfurt Defender reasons with the incompatibilist as follows:

> Let us work with whatever brand of libertarianism you want (noncausal, agent causal, or even indeterministic). With one exception, I will grant all its constituents to you, *including the formal incompatibilist requirement of indeterminism.* The exception is just the very point I want to contest—the necessity of alternative possibilities. On *that*, I request that you suspend judgment. Now let me introduce an example that should convince you that alternative possibilities are irrelevant to moral responsibility and the sort of freedom required for it.

Call this the *Broad Dialectical Situation.* On this interpretation of the situation, it should be clear that the Dilemma Defender's appeal to the first horn *is unimpeachable*; it is simply a condition of the debate that the Frankfurt Defender offer a case that does not presume determinism.

Why not couple a committed incompatibilist audience with a narrow interpretation of the first horn? Assuming that Frankfurt's argument is directed to a committed incompatibilist audience, it would be dialectically unfitting to assume that the first horn be assigned a narrow interpretation. If the Frankfurt Defender is going to grant libertarian freedom to an incompatibilist audience, and if the incompatibilist is going to allow Frankfurt a fair hearing, then the incompatibilist must be willing to put something on the table. She must be willing to concede that there is a subject of dispute which she is not simply entitled to treat in incompatibilist terms. If, in this dialectical context, what she would find objectionable about a Frankfurt example is the very fact that an agent in it had no alternative possibilities, then she would not in fact be prepared to suspend judgment on the very point Frankfurt wishes to dispute.

A further observation about the Broad Dialectical Situation merits attention. Suppose that the *compatibilist* Frankfurt Defender succeeds in convincing the incompatibilist that free and morally responsible agency does not require alternative possibilities.[13] The compatibilist Frankfurt Defender would *then* want to advance *further* arguments in opposition to the incompatibilist. This is because the sort of incompatibilist prepared to consider Frankfurt examples in this dialectical

---

[13] In discussing the Broad Dialectical Situation, we are fixing upon the *compatibilist* Frankfurt Defender. This, however, is only to bring into relief the further debates that might arise between compatibilists and incompatibilists if Frankfurt's argument is successful. We recognize that Frankfurt's argument is also available to incompatibilists who do not think that the incompatibility of determinism and free will turns upon alternative possibilities.

context would *not* be willing to concede that if Frankfurt's argument is sound, incompatibilism is *therefore* false.[14] Hence, in the Broad Dialectical Situation, this sort of incompatibilist, to begin with, approaches the debate over the Frankfurt-type examples under the assumption that perhaps determinism could be shown to rule out freedom and moral responsibility *independently* of ruling out alternative possibilities. And while this sort of incompatibilist might have initially assumed, prior to entertaining Frankfurt's argument, that it is sufficient to argue for incompatibilism by arguing that determinism rules out alternative possibilities, she would not be willing to grant that it is necessary. Were she to think it *necessary* to argue for incompatibilism by fixing upon alternative possibilities, then it would not make sense for her to suspend her commitment to alternative possibilities while considering a case of a libertarian free agent acting in a Frankfurt example.

Now consider a narrow interpretation of the first horn. As we argued above, the narrow interpretation of the first horn is not dialectically suited for a context in which Frankfurt's argument is meant to engage directly a committed incompatibilist audience. We therefore propose that the narrow interpretation is better suited for a context in which Frankfurt's argument is directed at those who are *undecided* about the free will problem, those who have not yet committed to compatibilism or incompatibilism. Indeed, irrespective of the compatibility question, this undecided audience is not even prepared to specify what sorts of constituents are necessary for the freedom moral responsibility requires. Hence, they have no settled convictions as to whether, for example, freely willed conduct requires higher-order identification with one's own springs of action, or if it requires responsiveness to reason, or if it requires alternative possibilities, and so forth. This audience amounts to what in the political landscape might be called "undecided voters."

Clearly, given *this* dialectical target—the undecided—a critic of Frankfurt's argument is *certainly* not entitled to the broad interpretation of the first horn. That is, the Dilemma Defender is not entitled to the complaint that the prior-sign in a Frankfurt example assumes a deterministic relation *simpliciter*. In this dialectical context, it is not to be assumed that this is objectionable; it might not be. On the other

---

[14] In fact, several prominent incompatibilists are in precisely this position. See, for example, Hunt; Pereboom, "Determinism *al Dente*," *Noûs*, XXIX (1995): 21–45, and *Living without Free Will*; Stump, "Libertarian Freedom and the Principle of Alternative Possibilities"; and Linda Zagzebski, "Does Libertarian Freedom Require Alternative Possibilities?" *Philosophical Perspectives*, XIV (2000): 231–48.

hand, relying upon the narrow interpretation, the Dilemma Defender might think it reasonable to object that if determinism already rules out alternative possibilities, then there is no work for a Frankfurt example to do. The counterfactual intervener adds nothing; the argument evaporates. We shall now develop this point.

Given a narrow interpretation of the first horn, consider the following dialectical context: the Frankfurt Defender invites an undecided audience to imagine a case of an agent who, prima facie, performs an uncontroversially blameworthy action, an action for which she is, prima facie, morally responsible. This audience makes no theoretical demands upon the metaphysical nature of this agent but simply assumes for argument's sake that we have a normally functioning person. Furthermore, no assumption is made as to whether the agent is determined, whether compatibilism or incompatibilism is true, or anything of the sort. In fact, it remains open whether the thesis of hard determinism is true and, hence, whether any agent acts of her own free will or is morally responsible for anything. (In this case, agents would merely be prima facie morally responsible for their *seemingly* blameworthy actions.) The Frankfurt Defender then asks this audience to accept that an agent's freedom and responsibility, if indeed any agent possesses these things, does not require alternative possibilities. So the Frankfurt Defender adds a counterfactual intervener to the case to do the work of eliminating the disputed alternatives, inviting the undecided audience to see that the lack of alternatives is *thereby* shown to be irrelevant to the blameworthiness of an agent's conduct. Now, given this dialectical context, the Dilemma Defender might protest:

> In a case in which the agent is determined, on one (admittedly disputed) view, the elimination of alternatives is rendered complete prior to the addition of the counterfactual machinery. And if an undecided audience grants that such an agent is even prima facie free and morally responsible without that machinery, as the incompatibilists might see it, this audience will have been duped into granting Frankfurt's point at the outset. Hence, a condition of this debate is that, absent Frankfurt's counterfactual machinery, it must be beyond dispute that an agent does have alternative possibilities.

Call this the *Narrow Dialectical Situation.* What is crucial to this dialectical situation is that the disputed ground is over the reasoned judgment of the theoretically uncommitted. If the Frankfurt Defender can win their allegiance, then she can show that the traditional incompatibilists (and also the traditional compatibilists) have been wrong all along to presume that basic intuitions about free will and moral

responsibility commandingly suggest that a freely willing agent must be able to do otherwise. Hence, the burden of proof would be shifted to those who asserted that it must.

In the Narrow Dialectical Situation, the Dilemma Defender's objection to the first horn is not without merit. However, it is unclear that it should amount to the charge that her opponent is begging the question. A Frankfurt Defender might insist that, dialectically, she is *not* inviting an undecided audience to imagine an agent who *does* act freely and is morally responsible, or who, absent the counterfactual intervener, *does* have alternative possibilities; she is only asking her undecided audience to acknowledge that these matters also require suspending judgment. Then the Frankfurt Defender can go to work on the Dilemma Defender's assumption that one manner of ruling out alternative possibilities (allegedly the assumption of determinism) renders other manners irrelevant to our understanding the conditions for free and morally responsible agency.

III. TWO TELLING RESPONSES TO THE DILEMMA DEFENSE

To appreciate the significance of the two dialectical situations, it will be helpful to consider two recent responses to the Dilemma Defense. Alfred Mele explains that, on a standard interpretation, the alternative possibilities at issue in PAP are "alternative *actions*, including intentional refrainings and instances of deciding and choosing."[15] Further, these possibilities are "actions like stealing from a poor-box and deciding not to steal from a poor-box, as opposed to more refined actions like '*S*'s stealing from a poor-box without having been made to do so by someone else' and '*S*'s deciding not to steal from a poor-box because Black made *S* decide not to do so'." Mele remarks:

> Now, if *PAP*...on the standard interpretation of alternative possibilities ...[is] falsified by some Frankfurt-style cases, then a thief...need not ever have been able to choose otherwise than he did at any point in his life, on the standard interpretation of choosing otherwise, in order to be morally responsible for his present choice and action. But notice that this is not to say that responsibility is compatible with *determinism*. For a traditional incompatibilist about determinism and moral responsibility, an agent's being deterministically caused to *A* suffices for his not being morally responsible for *A*-ing. So if Frankfurt-style cases are to persuade traditional incompatibilists that agents can be morally responsible for stealing from a poor-box, say, even though they could not have done otherwise, it must be a feature of these examples that the agents' stealing

---

[15] "Ultimate Responsibility and Dumb Luck," *Social Philosophy and Policy*, XVI (1999): 274–93, here p. 283.

> was not deterministically caused. In principle, a libertarian's incompatibilism might be motivated, not by the thought that determinism precludes our ever having been able to do otherwise than we did, but instead by the thought that in a deterministic world our actions...are ultimately causally ensured *consequences* of the laws of nature and states that obtained long before we were born (*ibid.*, p. 284).

The phrase in the passage, "So if Frankfurt-style cases are to persuade traditional Incompatibilists," bears noting. The phrase, along with Mele's insightful comments, provides convincing support that he means to be operating within the Broad Dialectical Situation.

In a co-authored piece with David Robb,[16] Mele's intriguing response to the Dilemma Defense modifies the Frankfurt-type example so that the fail-safe condition is not dependent upon any prior-sign. The condition triggers the desired action only in the counterfactual event that the agent does not freely perform the action. In fact, though, the agent performs the mental action of making a decision on her own without its being causally determined. The example includes two independent processes, each aimed at the same outcome, the agent's making some decision, such as the decision to rob a poor-box. One process is deterministic and will bring about the outcome under all circumstances save the one in which the other process brings about the outcome. The second is indeterministic—the agent's allegedly freely making the decision to rob the poor-box. As things actually unfold, the agent makes the decision herself. But if the indeterministic process were to fail to result in this outcome, the deterministic process would ensure that it occurs. The indeterministic process trumps the deterministic process *at* the moment of action. The agent's action is assumed to be a libertarian free one, but given the deterministic causal process that is trumped, the agent has no alternatives.

For our concerns, we note that it is partly but crucially in virtue of presupposing the Broad Dialectical Situation that the Mele/Robb response seems apt. (Whether it is convincing is another matter.[17]) Mele and Robb's modified no-prior-sign Frankfurt-type example is offered in the spirit of respecting their presumed interlocutor's concern that determinism may be inimical to free action and responsibility independently of expunging alternative possibilities. Hence, their approach makes no effort to challenge the first horn. As the Broad

---

[16] See Mele and Robb, "Rescuing Frankfurt-Style Cases," and also "BBs, Magnets, and Seesaws: The Metaphysics of Frankfurt-Style Cases."

[17] See, for example, Ginet's 2001 addendum to his "In Defense of the Principle of Alternative Possibilities," pp. 85–87.

Dialectical Situation dictates, they take the Dilemma Defender's first horn as unimpeachable.

Consider, in contrast, a strikingly different approach to the first horn. John Martin Fischer writes that it would not be "dialectically kosher" to assume that a determined agent in a Frankfurt example is morally responsible.[18] However, Fischer requests that it remain an open question. Fischer's more modest suggestion is merely that the presence of Frankfurt's intervener would not by itself be sufficient to show that the agent is *not* morally responsible. Hence, if the agent in a Frankfurt example is not morally responsible, it is not due to the agent's having alternative possibilities ruled out *by way of the counterfactual intervener.* This, as Fischer cautiously put it, "*suggests* that if Jones [an agent in a Frankfurt example] had no alternative possibilities at all [due to the truth of causal determinism], this would be irrelevant to the grounding of his moral responsibility."[19] So Fischer's response to the first horn involves three steps. First, he acknowledges that the Frankfurt Defender is not entitled to the assumption that a determined agent in a Frankfurt example is free and morally responsible. Second, he calls attention to the irrelevance of the counterfactual apparatus. Its presence alone would not explain why an agent in a Frankfurt example is not responsible (if she is not). Third, he holds that the irrelevance of eliminating alternative possibilities by way of the counterfactual intervener merely *suggests* that other means of ruling out alternative possibilities are also irrelevant.

Suppose that the Dilemma Defender presses Fischer by insisting that, in a deterministic scenario, he (Fischer) has not imagined a case in which it is clearly the counterfactual intervener that is ruling out the agent's alternative possibilities. If the agent is determined, then, the Dilemma Defender might object, *that* is why the agent does not possess alternative possibilities. And it would not be dialectically kosher to conclude that alternative possibilities are irrelevant in this context.[20] Hence, to insure that an important question is not begged, the Dilemma Defender might continue, we need a case in which the *only* manner of ruling out alternative possibilities is by way of the counterfactual intervener.

We acknowledge that the Dilemma Defender has a point. The Frankfurt Defender's case would be more convincing—there would

---

[18] "Free Will and Moral Responsibility," *Journal of Ethics* (forthcoming). See also "Recent Work on Moral Responsibility," and "Frankfurt-Type Examples and Semi-Compatibilism."

[19] "Free Will and Moral Responsibility."

[20] Stewart Goetz has pressed Fischer in just this way—see his "Frankfurt-Style Counterexamples and Begging the Question," *Midwest Studies in Philosophy*, XXIX (forthcoming 2005). See Fischer's reply in "Free Will and Moral Responsibility."

be less to contest—if it were beyond dispute that an agent's alternatives were ruled out *only* by virtue of the counterfactual intervener and not also by virtue of the truth of determinism. But is the Dilemma Defender entitled to claim that, in a case that does assume determinism, the Frankfurt Defender's argument rises to the level of begging the question? We believe that this turns upon the following assumption of the Dilemma Defender: if determinism makes it the case that an agent in a Frankfurt example cannot do otherwise, then the presence of the counterfactual intervener cannot aid in understanding the disputed condition for moral responsibility. Although we do not wish to evaluate the veracity of this assumption here, we agree with Fischer that it does not seem unreasonable—at least it is not question begging—for the Frankfurt Defender to reply to it as follows: the counterfactual intervener's presence in a deterministic scenario, while perhaps not necessary for ruling out alternative possibilities, is sufficient for doing so. Furthermore, its presence *is* irrelevant to any judgment that the agent is not free or morally responsible.[21] So why can this not help us to think about other factors that are alleged to rule out an agent's freedom or responsibility?[22]

Again, for our interests, we emphasize that Fischer's response to the first horn of the Dilemma Defense is designed to address a Narrow

---

[21] Some might insist that the presumption of the (alleged) irrelevance of the intervener in the case of a determined agent is weaker than in the case of the undetermined agent. In the case of the undetermined agent, assuming the intervener does rule out all robust alternatives, there will, minimally, be nonrobust alternatives (or morally insignificant alternatives), and these might be entailed by other conditions that are significant and necessary for moral responsibility (such as ultimacy). But this would be to shift the focus of the Frankfurt debate as most understand it. For a discussion of Frankfurt's argument in light of nonrobust alternatives, see Haji, "Flickers of Freedom, Obligation, and Responsibility," *American Philosophical Quarterly*, XL (2003): 287–302.

[22] In light of this discussion, consider Fischer's remarks:

In the Frankfurt-type scenario [assuming a determined agent], *two* causes make it the case that Jones [the agent] is unable to choose otherwise at [a certain time] *T2*: the prior condition of world (together with the laws of nature) and Black's counterfactual intervention. What the examples show is that *the mere fact* that Jones is unable to choose otherwise does not *in itself* establish that Jones is not morally responsible for his choice. This is because Black's counterfactual intervention is one of the things that make it the case that Jones is unable to choose otherwise at *T2*, and yet it is irrelevant to the grounding of Jones's moral responsibility. Considering this factor (the counterfactual intervention), and bracketing any other factor that might make it the case that Jones is unable to choose otherwise at *T2*, it seems to me that Jones may well be morally responsible for his action. The mere fact that he lacks alternative possibilities, then, cannot in itself be the reason that Jones is not morally responsible, if indeed he is not morally responsible—"Free Will and Moral Responsibility," Section IV.5.

Dialectical Situation. His response cannot plausibly be aimed at a direct attempt to convince the committed incompatibilist to give up the demand for alternative possibilities. It would not be reasonable to ask this audience to consider a *determined* agent who is a candidate for free and morally responsible agency. Rather, Fischer's response is meant to sway those who have not settled on whether determinism rules free and morally responsible agency. Furthermore, his response clearly presupposes a narrow interpretation of the first horn. He means to show that one manner of ruling out alternative possibilities does not impugn a philosophical lesson we can draw from another manner of ruling them out.

IV. DIALECTICAL DELICACIES AND THE CLAIM TO CONTESTED GROUND

We have been exploring hitherto unnoticed dialectical delicacies within the debate over the status of Frankfurt's argument against PAP.[23] Our discussion has helped distinguish two very different dialectical situations, complete with distinct argumentative burdens and ground rules as to what can be assumed either in advancing Frankfurt's argument or, instead, in objecting to it.

Under the assumption that Frankfurt's argument is meant to address directly a committed incompatibilist audience, it *is* dialectically required that the Frankfurt Defender proceed by offering an example that does not assume that the agent in it is determined. This is a burden that the Frankfurt Defender must accept in this dialectical context. But in this context, the incompatibilist who is prepared to entertain Frankfurt's argument is not entitled to demand that (absent Frankfurt's counterfactual machinery) the relevant case involves an agent with alternative possibilities. Demanding this is tantamount to refusing Frankfurt a fair hearing. Hence, fitting Frankfurt's argument for a committed incompatibilist audience requires a Broad Dialectical Situation. In the Broad Dialectical Situation, what the Dilemma Defender finds objectionable about the deterministic horn of the posed dilemma is that determinism obtains simpliciter. Given the dialectical terms in which the Frankfurt Defender enters into *this* debate, we agree that the Dilemma Defender's claim to the deterministic horn is beyond dispute. It simply makes no sense for the Frankfurt Defender to consider resistance on this front. Hence, in this context, the Frank-

---

[23] To our knowledge, the only other discussion of varying dialectical assumptions is in Pereboom's recent article, "Source Incompatibilism and Alternative Possibilities," see pp. 190–93. Some of Pereboom's observations are similar to ours. However, he offers a different diagnosis of the dialectical forces at work in the debate.

furt Defender's only plausible option is to resist the Dilemma Defense on its other horn, the horn in which it is assumed that the agent is not determined. (This is precisely the strategy adopted by Mele.)

But we also wish to note that those incompatibilists who insist upon the Broad Dialectical Situation—requiring that Frankfurt begin with a case of a libertarian free agent—are dialectically committed at least to the possibility that determinism might rule out free will and moral responsibility for reasons other than ruling out alternative possibilities. If so, certain sorts of incompatibilists are not entitled to argue with Frankfurt within the parameters of the Broad Dialectical Situation. Some incompatibilists hold that arguments for incompatibilism *require* a defense of the incompatibility of determinism and alternative possibilities. They believe that the only reason that determinism rules out free will and moral responsibility is because it rules out alternative possibilities. If these incompatibilists wish to grapple with Frankfurt given a Broad Dialectical Situation, we believe that they do so in bad faith. By demanding that Frankfurt begin with a case of a libertarian free agent, they are demanding that Frankfurt provide a case in which any freely willing and morally responsible agent does have alternative possibilities. This of course is just what Frankfurt wants to disprove.

Hence, we challenge incompatibilists to consider their reasons for incompatibilism. If their only reasons have to do with alternative possibilities, they must come clean. In doing so we believe that they must concede that they are not entitled to the Dilemma Defense if it is meant to address a Broad Dialectical Situation. If these incompatibilists wish to rely upon the Dilemma Defense, then they must accept the conditions of the debate established by the Narrow Dialectical Context.[24]

---

[24] This challenge is especially poignant for Kane. He is hard to fit within the dialectical parties as we have sorted them out. On the one hand, Kane maintains that an alternative possibilities condition, an AP condition (a PAP-like condition), is only one of two conditions necessary for libertarian free will and moral responsibility. According to Kane, another is an ultimacy condition. Ultimacy requires that an agent's freely willed actions ultimately trace back to factors that originate in the agent and cannot be traced back to causally sufficient conditions obtaining independently of the agent (*The Significance of Free Will*, pp. 58–59). Hence, it seems that Kane is a theorist who can insist upon the Broad Dialectical Situation. He can argue that there are grounds—the grounds of ultimacy—independent from considerations of alternative possibilities that could be used to establish the incompatibility of free will and determinism. However, Kane argues that ultimacy itself requires dual control (control over different courses of action) and that dual control entails alternative possibilities—"The Dual Regress of Free Will and the Role of Alternative Possibilities," *Philosophical Perspectives*, XIV (2000): 57–79. On our view of the dialectical terrain, if Kane wishes to do battle with the Frankfurt Defender within the Broad Dialectical Context, then he should reconsider his claim that ultimacy entails alternatives. If he is unwilling to concede this point, we believe that he is forced to argue with the Frankfurt Defender in a Narrow Dialectical Context.

In the narrow context, the presumed target of Frankfurt's argument is the theoretically uncommitted voter. Thus, the Frankfurt Defender does not have to initiate the debate by beginning with the case of a libertarian free agent. In this context, suggesting that an agent in a Frankfurt example is determined (simpliciter) is not necessarily objectionable, and the Dilemma Defender is not entitled to this as a complaint. However, as we have argued, the Dilemma Defender might understandably object in this narrow context that a determined agent's alternatives are ruled out prior to the addition of Frankfurt's counterfactual machinery. So the Dilemma Defender *can* reasonably argue that the first horn of the dilemma objects to a determined agent *because* such agents (allegedly) lack alternative possibilities. However, we have also argued that the Dilemma Defender's objection, when understood in the context, is at least disputable. In this case, the Dilemma Defender does not have a clear claim to the charge that the Frankfurt Defender is begging the question. In this context, the Frankfurt Defender has a plausible reply: even if determinism rules out alternative possibilities, so does the presence of the Frankfurt's counterfactual machinery. The latter way of ruling out alternatives is irrelevant to an evaluation of an agent's freedom and responsibility, and this provides plausible grounds for the suggestion that the former is as well. Our concern here is not to settle this dispute, only to point out that, dialectically, in this context, the Dilemma Defender cannot claim clear victory with respect to the first horn. On this horn, in this context, her charge of begging the question is unfounded. There is a legitimate dispute here, and the Frankfurt Defender deserves a hearing. (This is precisely the strategy adopted by Fischer.)

So in one dialectical context, the Broad one, the Dilemma Defender is entitled to the complaint that the first horn of the dilemma is question begging. In another dialectical context, the Narrow one, the Dilemma Defender is still entitled to the complaint that on the first horn the agent is determined, but her complaint does not amount to the devastating charge that the Frankfurt Defender is begging the question. We have argued that only certain sorts of incompatibilists are entitled to the Broad Dialectical Context, namely those who acknowledge that determinism might undermine free will and determinism for reasons other than ruling out alternative possibilities. Incompatibilists unwilling to acknowledge this will have to operate within the Narrow Dialectical Situation, and at least with respect to the first horn of the Dilemma Defense, they have their work cut out for them. In this dialectical context, the Frankfurt Defender, such as Fischer, has a dialectically fair and reasonable reply.

V. CONCLUSION

In conclusion, differentiating between two dialectical situations, it is natural to reflect on the big question: What is the appropriate dialectical situation that the Frankfurt examples should be thought to address, the Broad or the Narrow? Of course, argumentative strategies can be put to different sorts of purposes, and we do not wish to suggest that there is one "correct" context that Frankfurt's argument should be taken to address. On the one hand, the Narrow Dialectical Situation will not likely impress a committed incompatibilist. The committed incompatibilist will likely object that an argument merely designed to win the hearts of the uncommitted voter cannot speak to a theorist with deeper libertarian commitments regarding the conditions of agency.[25] Hence, the Frankfurt Defender wishing to address directly this sort of incompatibilist should be prepared to work within the framework of the Broad Dialectical Situation. In this case, the Frankfurt Defender should accept that she will only be able to engage the incompatibilist if she begins a dialogue by doing battle on the libertarian's own turf. On the other hand, why should the disputed turf between compatibilists and incompatibilists be granted to incompatibilists in the first place—as if it is the compatibilists' burden to win it back? The Frankfurt Defender who initiates the dialectic within the Broad Dialectical Situation has already conceded that the compatibilists have the burden of proof—*they* need to overcome the intuitive presumption in favor of both the demand for alternative possibilities and an incompatibilist understanding of them. But this seems to initiate the debate in the wrong place. Frankfurt-type examples, it seems to us, are best understood as ones that question the pre-theoretical intuitions that inform one's theorizing about the issues of free action and moral responsibility. They are, in short, thought experiments to settle, first, what turf should be the subject of controversy between compatibilists and incompatibilists. *If* this is so, then the Narrow Dialectical Context is the appropriate one for understanding the force of Frankfurt's argument against the importance of alternative possibilities for free will and moral responsibility.

                                                   ISHTIYAQUE HAJI
University of Calgary

                                                   MICHAEL MCKENNA
Ithaca College

---

[25] For example, in *The Significance of Free Will*, Kane maintains the considerations of ultimacy are also needed for free and morally responsible agency and that ultimacy is incompatible with determinism.

COMMENTS AND CRITICISM

STOPPING TO REFLECT*

Our note is prompted by a recent article by Frank Arntzenius, "Some Problems for Conditionalization and Reflection."[1] Through a sequence of examples, that article purports to show limitations for a combination of two inductive principles that relate current and future rational degrees of belief: *Temporal Conditionalization* and *Reflection*:

(i) *Temporal Conditionalization* is the rule that, when a rational agent's corpus of knowledge changes between *now* and *later* solely by learning the (new) evidence, *B*, then coherent degrees of belief are updated using conditional probability according to the formula, for each event *A*,

$$P_{later}(A) \, = \, P_{later}(A \, | \, B) \, = \, P_{now}(A \, | \, B).$$

(ii) *Reflection*[2] between *now* and *later* is the rule that current conditional degrees of belief defer to future ones according to the formula that, for each event *A*,

$$P_{now}(A \, | \, P_{later}(A) \, = \, r) \, = \, r.[3]$$

We will use the expression "Reflection holds with respect to the event *A*" to apply to this equality for a specific event *A*.

It is our view that neither of these principles is mandatory for a rational agent.[4] However, we do not agree with Arntzenius that, in the examples in his article, either of these two is subject to new

[1] This JOURNAL, C, 7 (July 2003): 356–70.
[2] See Bas C. van Fraassen's "Belief and the Will," this JOURNAL, LXXXI, 5 (May 1984): 235–56. van Fraassen's *Reflection* has an antecedent in Michael Goldstein's "Prevision of a Prevision," *JASA*, LXXVIII (1983): 817–19.
[3] Here and through the rest of this note '*r*' is a standard designator for a real number—this in order to avoid a Miller-styled problem—see David Miller's "A Paradox of Information," *British Journal for the Philosophy of Science*, XVII (1966): 59–61.
[4] We have argued, for example, that when (subjective) probability is finitely but not countably additive, then there are simple problems where (i) is reasonable, but where (i) precludes (ii). See our "Reasoning to a Foregone Conclusion," *JASA*, XCI (1996): 1228–36. Also, Isaac Levi argues successfully, we think, that (i) is not an unconditional requirement for a rational agent—see his "The Demons of Decision," *The Monist*, LXX (1987): 193–211.

restrictions or limitations beyond what is already assumed as familiar in problems of stochastic prediction.

To the extent that a rational person does not know *now* exactly what she or he will know in the future, anticipating one's future beliefs involves predicting the outcome of a stochastic process. The literature on stochastic prediction relies on two additional assumptions regarding states of information and the temporal variables that index them[5]:

(iii) When $t_2 > t_1$ are two fixed times, then the information the agent has at t_2 includes all the information that she or he had at time t_1.[6] This is expressed mathematically by requiring that the collection of information sets at all times through the future form what is called a *filtration*.

Second, since the agent may not know *now* the precise time at which some specific information may become known in the future, then future times are treated as *stopping times*. That is:

(iv) For each time T (random or otherwise) when a prediction is to be made, the truth or falsity of the event $\{T \leq t\}$ is known at time t, for all fixed t. Such (random) times T are called *stopping times*.

In this note, we apply the following three results[7] to the examples in Arntzenius's article. These results, we believe, help to explain why the examples at first appear puzzling and why they do not challenge either Temporal Conditionalization or Reflection. Result 1 covers the ordinary case, where Reflection holds. Results 2 and 3 establish that Reflection will fail when one or the other of the two additional assumptions, (iii) and (iv), fail. It is not hard to locate where these assumptions are violated in the examples that Arntzenius presents.

Result 1. When "*later*" is a *stopping time*, when the information sets of future times form a *filtration*, and assuming that degrees of belief are updated by Temporal Conditionalization, then Reflection between *now* and *later* follows.

Result 2. When the information known to the agent over time *fails*

---

[5] See, for example, Patrick Billingsley, *Probability and Measure*, 3rd edition (New York: Wiley, 1995), section 35.

[6] Here and through the rest of this note, 't' is a standard designator for a real number for time. More precisely, we use the subscripted variable, for example 't_1' to denote a specific time as the agent of the problem is able to measure it. We presume that the agent has some real-valued "clock" that quantifies a transitive relation of "is later than." Subtleties about the differences between how time is so indexed for different observers is relevant to one of Arntzenius's puzzles, to wit, the Prisoner's Problem.

[7] Proofs for these three are given in the appendix. In this note, we assume that all probability is countably additive.

to form a filtration, not only is Temporal Conditionalization vacuously satisfied (as its antecedent fails), but then Reflection fails unless what is forgotten in the failure of filtration becomes practically certain (its probability becomes 0 or 1) in time for future predictions, *later*.

Result 3. However, if the information known to the agent over time forms a filtration and Temporal Conditionalization holds, but "*later*" is not a *stopping time*, then Reflection between *now* and *later* holds for the specific event A, that is, $P_{now}(A \mid P_{later}(A) = r) = r$, subject to the necessary and sufficient condition, (3.1), below.

Let H_t be the event "$t = later$." When *later* is not a stopping time, the event H_t is news to the agent making the forecasts. The question at hand is whether this news is relevant to the forecasts expressed by Reflection. To answer that question, concerning such forecasts about the event A, define the quantity $y_t(A)$ by

$$y_t(A) = \frac{P_{now}(H_t \mid P_t(A) = r \ \& \ A)}{P_{now}(H_t \mid P_t(A) = r)}.$$

The quantity $y_t(A)$ is an index of the current conditional dependence between A and H_t, given that $P_t(A) = r$. For example, $y_t(A) = 1$ if and only if A and H_t are conditionally independent for the agent, *now*, given that $P_t(A) = r$. In other words, by symmetry of conditional independence, $y_t(A) = 1$ if and only if the agent's current conditional probability of A given that $P_{later}(A) = r$, is unchanged by the added information H_t.

Reflection holds for A between *now* and *later*, $P_{now}(A \mid P_{later}(A) = r) = r$ if and only if, given $P_{later}(A) = r$, the conditional expected value $y_T(A) = 1$. Specifically, if and only if

(3.1) $1 = \Sigma_t \ y_t(A) \ P_{now}(H_t \mid P_{later}(A) = r)$

Thus, Reflection is satisfied between *now* and *later* if and only if (3.1) holds for each A.

Next, we illustrate the second and third results with examples that show how Reflection may fail.

Example 1 (illustrating Result 2). Suppose that the agent will observe a sequence of coin tosses, one at a time at a known rate, for example one toss per minute. Let $X_n = 1$ if the coin lands heads up on toss n, and let $X_n = 0$ otherwise. The agent does not know how the coin is loaded, but believes that it is *fair* (event A) with personal probability $1/2$, and that with personal probability $1/2$ it is biased with a chance of $3/4$ for landing tails (event A^c). Also, he believes that tosses are conditionally independent given the loading, that is, given that the coin is *fair* or given that it is biased $3/4$ for tails.

Time is indexed for the agent by the number of the most recent coin toss. The time "*now*" occurs after the first toss ($t = n = 1$), and "*later*" denotes the time ($t = n = 2$) just after the second toss. Unfortunately, at each time t, the agent knows that he can remember only the most recent flip, X_t, though he knows which numbered toss it is because, for example, he can see a clock. Suppose that the first toss lands heads up, which is the event $C = \{X_1 = 1\}$. The information that will be available to the forgetful agent later (at $t = 2$) will be only that either $B_1 = \{X_2 = 1\}$ or $B_0 = \{X_2 = 0\}$. He will not recall C because of his predictable memory lapse, and he knows all this. It is straightforward to compute:

$$P_{later}(A \,|\, B_1) \;=\; 2/3 \text{ and } P_{later}(A \,|\, B_0) \;=\; 2/5.$$

However, at $t = 1$, the agent's conditional probability for A, given event B_1 occurring at $t = 2$, satisfies $P_{now}(A \,|\, B_1) = 4/5$. Similarly, if *now* he conditions on event B_0 occurring at $t = 2$, his conditional probability will satisfy $P_{now}(A \,|\, B_0) = 4/7$.

Of course, Temporal Conditionalization holds vacuously at the *later* time, since the information sets available to the agent do not form a filtration. Reflection fails in this setting, as the agent does not remember at the *later* time what happened *now*, and he knows this all along. If B_1 occurs then $P_{later}(A) = P_{later}(A \,|\, B_1) = 2/3$, and if B_0 occurs then $P_{later}(A) = P_{later}(A \,|\, B_0) = 2/5$. Hence,

$$P_{now}(A \,|\, P_{later}(A) = 2/3) \;=\; 4/5$$

and

$$P_{now}(A \,|\, P_{later}(A) = 2/5) \;=\; 4/7.$$

Example 2 (illustrating Result 3 when condition (3.1) fails and then Reflection fails too). Modify Example 1 so that the agent has no memory failures and updates his degrees of belief by Temporal Conditionalization. Also, change the time "*now*" to denote the minute prior to the first toss, that is, *now* is $t = n = 0$. Define the time "*later*" to be the *random time, T, just prior to the first toss that lands heads up.* From the point of view of the agent, the quantity T is not an observable random variable up to and including time T, and it is not a *stopping time* either. It is observable to the agent starting with time $T + 1$, of course, as by then he will have seen when the first head occurs.

With probability 1 the possible values for T are $T = 0, 1, 2,\dots$. It is straightforward to verify that: $P_{later}(A) = [1 + (3/2)^n]^{-1}$, when $T = n$, for $n = 0, 1, 2,\dots$. Notice that $P_{later}(A) \le 1/2$, no matter when T occurs, and $P_{later}(A) < 1/2$ for $T > 0$, since if $T > 0$, the initial sequence of tosses that the agent observes all land tails up. However, from the

value of $P_{later}(A)$ and knowing it is this quantity, one may calculate T exactly and thus know the outcome of the $n+1^{st}$ toss, which is heads. But when the agent computes $P_{later}(A)$ at the time *later*, he does not then know that *later* has arrived. Thus, *later*, he is not in a position to use the extra information that he would get from knowing when T occurs to learn the outcome of the $n+1^{st}$ toss. To repeat the central point, T is not a stopping variable.

It is evident that Reflection fails, $P_{now}(A | P_{later}(A) = r) \neq P_{later}(A)$. The extra information, namely that $P_{later}(A) = r$ rather than merely that $P_t(A) = r$ where t is the time on the agent's clock, is information that is relevant to his *current* probability of A, since it reveals the outcome of the next toss. Even *now*, prior to any coin tosses, when he computes $P_{now}(A | P_{later}(A) = r)$, the conditioning event reveals to him the value of T, since n is a function of r. In this case, the conditioning event entails the information of n and when the first heads occurs, namely, on the $n+1^{st}$ toss. Then Reflection fails as

$$P_{now}(A | P_{later}(A) = [1 + (3/2)^n]^{-1}) = (1 + 3^n/2^{n+1})^{-1}.$$

It remains only to see that (3.1) fails as well. Consider the quantity $y_t(A)$ used in condition (3.1). $y_t(A) = \dfrac{P_{now}(H_t | P_t(A) = r \& A)}{P_{now}(H_t | P_t(A) = r)}$. Given $P_t(A) = r$, the added information that A obtains is relevant to the agent's current probability when *later* occurs. Specifically, as $P_t(A) = [1 + (3/2)^n]^{-1}$ entails that $t = n$, $P_{now}(H_t | P_t(A) = [1 + (3/2)^n]^{-1}) = P_{now}(X_{t+1} = 1 | P_t(A) = [1 + (3/2)^n]^{-1}) = (1/2)[1+(3/2)^n]^{-1} + (1/4)(3/2)^n [1+(3/2)^n]^{-1} < 1/2 = P_{now}(X_{t+1} = 1 | P_t(A) = [1+(3/2)^n]^{-1} \& A) = P_{now}(H_t | P_t(A) = [1 + (3/2)^n]^{-1} \& A)$. Thus, $y_t > 1$. Hence, $1 < \Sigma_t y_t(A) P_{now}(H_t | P_{later}(A) = r)$.

Example 3 (illustrating Result 3 when (3.1) obtains and Reflection holds even though *later* is not a *stopping time*). In this example, consider a sequence of three times, $t = 0$, 1, and 2. *Now* is time $t = 0$. The available information increases with time, so that the information sets form a filtration, and the agent updates his degrees of belief by Temporal Conditionalization. Let the random time *later* be one of the two times $t = 1$, or $t = 2$, chosen at random, but which one is not revealed to the agent. Let the event H_i be that *later* $= i$, ($i = 1$, 2) and suppose that the occurrence of H_i (or its failure) while not known to the agent at any of the three times is independent of all else that the agent does know at all three times. In this case, for each event A (even for $A = H_i$) equation (3.1) is satisfied. That is, by the assumptions of the problem, either $y_t(A) = \dfrac{P_{now}(H_t | P_t(A) = r \& A)}{P_{now}(H_t | P_t(A) = r)} = 1$,

or if $A = H_i$ then $y_t(A) = \dfrac{P_{now}(H_t \mid P_{later}(A) = r)}{P_{now}(H_t \mid P_t(A) = r)} = 1$. Thus, $P_{now}(A \mid P_{later}(A) = r) = r$. That is, even though *later* is not a stopping time, Reflection holds in this case since, given that $P_{later}(A) = r$, no new (relevant) evidence about A is conveyed through knowing that *later* has arrived, H_i.

We note that Result 2 applies to the Sleeping Beauty,[8] Shangri La, and Duplication examples of Arntzenius's article, where known failures of memory are explicit to the puzzles. Result 3 applies to explain the failure of Reflection in the two versions of the "Prisoner" example where the *local time* in the story, as measured by an ordinary clock (for example, "11:30 PM" in John Collins's example) is not a *stopping time* for the Prisoner.

It is our impression of Collins's Prisoner example that the reader is easily mistaken into thinking that the *local time*, as measured by an ordinary clock in the story, is a *stopping time* for all the characters in the story. Then Reflection holds for each of them, in accord with Result 1. In Collins's example, the *local time*, for example, 11:30 PM, is a *stopping time* for the Jailor (and also for the reader), but *not* for the Prisoner. For the Prisoner, time is measured by real-valued increments over the starting point, denoted by "now." Increments of *local time* are *stopping times* for the Prisoner. This is because the Prisoner does not know at the start of the story which of two *local times* equals his time *now*. For subsequent times, he does know how much *local time* has elapsed since *now*. But that information is not equivalent to knowing the *local time*. That difference in what is a *stopping time* for different characters is what makes this a clever puzzle, we think.

<div align="center">APPENDIX</div>

Proof of Result 1.[9] Assume that when X is a random variable and C is an event, the agent's expected value $E_P(X)$ and conditional expected

---

    [8] See also J.Y. Halpern's "Sleeping Beauty Reconsidered: Conditioning and Reflection in Asynchronous Systems," Department of Computer Science, Cornell University (September 2003). We agree with Halpern that, in our words, coherence of a sequence of previsions does not require that they will be *well calibrated*—in a frequency sense of "well calibrated." That is, we think it is reasonable for Sleeping Beauty to give a prevision of $1/2$ to the event that the known fair coin landed heads on the flip in question, each time she is woken up. What complicates the analysis is that the repeated trials in Sleeping Beauty's game do not form an independent sequence, and her mandated forgetfulness precludes any "feedback" about the outcome of past previsions. When repeated trials are dependent and there is no learning about past previsions, coherent previsions may be very badly calibrated in the frequency sense. For other examples and related discussion of this point see, for example, Seidenfeld, "Calibration, Coherence, and Scoring Rules," *Philosophy of Science*, LII (1985): 274–94.
    [9] van Fraassen, "Belief and the Problem of Ulysses and the Sirens," *Philosophical*

value $E_P(X|C)$ exist with respect to the probability P. Let A be an event and let $X = P(A|Y)$ be a random variable, a function of the random variable Y. Then, as a consequence of the law of total probability, with C also a function of Y,

$$(1.1) \quad P(A|C) = E_P[X|C].$$

Assume that the agent's degrees of belief *now* include his *later* degrees of belief as objects of uncertainty. That is, future events such as "$P_{later}(A) = r$" and "$P_{later}(A|C) = q$" are proper subjects, *now*, of the agent's current degrees of belief. Suppose that, *now*, the agent anticipates using (i) Temporal Conditionalization in responding to the new evidence $Y = y$ that he knows he will learn at the stopping time, *later*. For example, Y might be the result of a meter reading made at the *later* time, with a sample space of m many possible values $Y = \{y_1,...,y_m\}$. Thus, by (i), for whichever value y of Y that results,

$$(1.2) \quad P_{later}(A) = P_{later}(A \,|\, Y = y) = P_{now}(A \,|\, Y = y).$$

Then, by (i) and (1.1), for C also a function of Y, the agent *now* believes that

$$(1.3) \quad P_{now}(A \,|\, C) = E_{P_{now}}[P_{later}(A) \,|\, C].$$

Let C be the event, "$P_{later}(A) = r$," which we presume is a possible value for $P_{later}(A)$ from the agent's current point of view. (This C is a function of Y.) Then, because *later* is a stopping time,

$$(1.4) \quad P_{now}(A \,|\, P_{later}(A) = r) = E_{P_{now}}[P_{later}(A) \,|\, P_{later}(A) = r)].$$

As

$$(1.5) \quad E_{P_{now}}[P_{later}(A) \,|\, P_{later}(A) = r)] = r,$$

therefore

$$(1.6) \quad P_{now}(A \,|\, P_{later}(A) = r) = r,$$

that is, then Reflection holds as well.

Proof of Result 2. To show that Reflection fails, consider two times $t_1 < t_2$. Call an event *forgotten* if its truth or falsity is known at time t_1 but not at time t_2. From the assumption that these times do not form a filtration, let E be forgotten between t_1 and t_2 and allow that at t_1

Studies, LXXVII (1995): 7–37, argues (pp. 17–19) that Temporal Conditionalization implies *Reflection*. His argument (pp. 18–19) has an additional, tacit assumption that the time t at which conditioning applies for *Reflection* is a stopping time.

this is known to happen at t_2. Since $P_{t_1}(E) \in \{0,1\}$, conditioning will not change this value, that is,

(2.1) $P_{t_1}(E) = P_{t_1}(E \mid P_{t_2}(E) = r)$

for a set of r-values of probability 1 under P_{t_1}. But, since it is known at t_1 that E will be forgotten at t_2, $P_{t_1}(0 < P_{t_2}(E) < 1) = 1$. Hence Reflection fails as $0 < r < 1$ in (2.1).

Proof of Result 3. Assume that the agent's information sets form a filtration over time and that Temporal Conditionalization holds between *now* and *later* but that *later* is not a stopping time for the agent. Let H_t be the event "*later* = t" for the specific time t. That is, assume that $0 < P_{later}(H_t) < 1$, when *later* occurs at t.

Later is the future time we will focus on in calculating whether Reflection holds, that is, we will inquire whether, for each event A, $P_{now}(A \mid P_{later}(A) = r) = r$, or not. We calculate as follows.

$P_{now}(A \mid P_{later}(A) = r)$

$\quad = \Sigma_t\, P_{now}(A\ \&\ H_t \mid P_{later}(A) = r)$

by the law of total probability.

$\quad = \Sigma_t\, P_{now}(A \mid P_{later}(A) = r\ \&\ H_t)\, P_{now}(H_t \mid P_{later}(A) = r)$

by the multiplication theorem

$\quad = \Sigma_t \dfrac{P_{now}(H_t \mid P_t(A) = r\ \&\ A)}{P_{now}(H_t \mid P_t(A) = r)}\, P_{now}(A \mid P_t(A) = r)\, P_{now}(H_t \mid P_{later}(A) = r)$

by Bayes's theorem and the equivalence of

$(P_{later}(A) = r\ \&\ H_t)$ and $(P_t(A) = r\ \&\ H_t)$

$\quad = r\, \Sigma_t \dfrac{P_{now}(H_t \mid P_t(A) = r\ \&\ A)}{P_{now}(H_t \mid P_t(A) = r)}\, P_{now}(H_t \mid P_{later}(A) = r)$

as $P_{now}(A \mid P_t(A) = r) = r$ by Result 1.

$\quad = r\, \Sigma_t\, y_t(A)\, P_{now}(H_t \mid P_{later}(A) = r).$

by the definition of $y_t(A)$

Hence, $P_{now}(A \mid P_{later}(A) = r) = r$ if and only if $\Sigma_t\, y_t(A)\, P_{now}(H_t \mid P_{later}(A) = r) = 1$, which is condition (3.1).

M.J. SCHERVISH, T. SEIDENFELD, J.B. KADANE
Carnegie Mellon University

NEW BOOKS

ALLEN, BARRY: *Knowledge and Civilization.* Boulder, CO: Westview, 2004. x, 342 p. Cloth $95.00, paper $35.00.

ALLISON, HENRY E.: *Kant's Transcendental Idealism: An Interpretation and Defense,* Revised and Enlarged Edition. New Haven: Yale, 2004. xix, 537 p. Paper $25.00.

ALTIERI, CHARLES: *The Particulars of Rapture: An Aesthetics of the Affects.* Ithaca, NY: Cornell, 2003. x, 299 p. Cloth $49.95, paper $22.50.

AMERIKS, KARL: *Interpreting Kant's "Critiques."* New York: Oxford, 2003. vii, 351 p. Cloth $74.00.

ARONSON, RONALD: *Camus and Sartre: The Story of a Friendship and the Quarrel that Ended It.* Chicago: University Press, 2004. x, 291 p. Cloth $32.50.

AZZOUNI, JODY: *Deflating Existential Consequence: A Case for Nominalism.* New York: Oxford, 2004. viii, 241 p. Cloth $49.95.

BAKER, GORDON: *Wittgenstein's Method: Neglected Aspects.* Malden, MA: Blackwell, 2004. xi, 312 p. Cloth $59.95.

BANKS, ERIK C.: *Ernst Mach's World Elements: A Study in Natural Philosophy.* Norwell, MA: Kluwer, 2003. xii, 292 p. Cloth $127.00.

BAUHN, PER: *The Value of Courage.* Lund, Sweden: Nordic, 2003. 213 p. Cloth $60.00.

BAUM, ERIC B.: *What Is Thought?* Cambridge: MIT, 2004. xi, 478 p. Cloth $40.00.

BEISER, FREDERICK C.: *The Romantic Imperative: The Concept of Early German Romanticism.* Cambridge: Harvard, 2003. xiii, 243 p. Cloth $45.00.

BEK, LISE: *Reality in the Mirror of Art.* Gylling, Denmark: Aarhus, 2004. 572 p. Cloth $65.95.

BELL, LINDA A.: *Beyond the Margins: Reflections of a Feminist Philosopher.* Albany: SUNY, 2003. x, 245 p. Cloth $59.50, paper $17.95.

BERTRAM, CHRISTOPHER: *Rousseau and the Social Contract.* New York: Routledge, 2004. x, 214 p. Paper $15.95.

BLOOM, PAUL: *Descartes' Baby: How the Science of Child Development Explains What Makes Us Human.* New York: Perseus, 2004. xv, 271 p. Cloth $26.00.

BODAN, MARGARET A.: *The Creative Mind: Myths and Mechanisms,* Second Edition. New York: Routledge, 2004. xiii, 344 p. Paper $21.95.

BOVENS, LUC, and STEPHAN HARTMANN: *Bayesian Epistemology.* New York: Oxford, 2003. ix, 159 p. Cloth $60.00, paper $27.50.

BRANN, EVA, with PETER KALKAVAGE, and ERIC SALEM: *The Music of the Republic: Essays on Socrates' Conversations and Plato's Writings.* Philadelphia: Paul Dry, 2004. xiv, 378 p. Cloth $24.95.

BRINK, DAVID O.: *Perfectionism and the Common Good: Themes in the Philosophy of T.H. Green.* New York: Oxford, 2003. xiv, 139 p. Cloth $27.50.

BROWN, JESSICA: *Anti-individualism and Knowledge.* Cambridge: MIT, 2004. xiv, 339 p. Paper $25.00.

BUNGE, MARIO: *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge.* Toronto: University Press, 2003. xiii, 330 p. Cloth $75.00.

CAMPBELL, JOSEPH: *The Hero with a Thousand Faces,* Second Edition. Princeton: University Press, 2004. lxvi, 403 p. Cloth $29.95.

CAPALDI, NICHOLAS: *John Stuart Mill: A Biography.* New York: Cambridge, 2004. xx, 436 p. Cloth $40.00.

CARRUTHERS, PETER: *The Nature of the Mind: An Introduction.* New York: Routledge, 2004. xii, 308 p. Paper $24.95.

CAVELL, STANLEY: *Cities of Words: Pedagogical Letters on a Register of the Moral Life.* Cambridge: Harvard, 2004. xi, 458 p. Cloth $29.95.

CHEAH, PHENG: *Spectral Nationality: Passages of Freedom from Kant to Postcolonial Literatures of Liberation.* New York: Columbia, 2003. xvi, 408 p. Cloth $62.50, paper $22.50.

CHIHARA, CHARLES S.: *A Structural Account of Mathematics.* New York: Oxford, 2004. xiv, 380 p. Cloth $70.00.

CLARK, DAVID K.: *Empirical Realism: Meaning and the Generative Foundation of Morality.* Lanham, MD: Lexington, 2004. ix, 419 p. Cloth $75.00.

CLARK, DESMOND M.: *Descartes's Theory of Mind.* New York: Oxford, 2003. viii, 267 p. Cloth $49.95.

CLARKE, RANDOLPH: *Libertarian Accounts of Free Will.* New York: Oxford, 2003. xv, 244 p. Cloth $49.95.

COOPER, DAVID E.: *Meaning.* Ithaca, NY: McGill-Queen's, 2004. 152 p. Cloth $70.00, paper $22.95.

COOPER, GREGORY J.: *The Science of the Struggle for Existence: On the Foundations of Ecology.* New York: Cambridge, 2003. xv, 319 p. Cloth $60.00.

CRAPANZANO, VINCENT: *Imaginative Horizons: An Essay in Literary-Philosophical Anthropology.* Chicago: University Press, 2004. xiii, 260 p. Cloth $50.00, paper $19.00.

CURTLER, HUGH MERCER: *Ethical Argument: Critical Thinking in Ethics,* Second Edition. New York: Oxford, 2004. xx, 184 p. Paper $19.95.

DA COSTA, NEWTON C. A., and STEVEN FRENCH: *Science and Partial Truth: A Unitary Approach to Models and Scientific Reasoning.* New York: Oxford, 2003. 259 p. Cloth $49.95.

DAVIES, BRIAN: *An Introduction to the Philosophy of Religion,* Third Edition. New York: Oxford, 2004. ix, 333 p. Paper $24.95.

DAVIES, DAVID: *Art as Performance.* Malden, MA: Blackwell, 2004. xii, 278 p. Paper $29.95.

DAY, JERRY: *Voegelin, Schelling, and the Philosophy of Historical Existence.* Columbia: Missouri UP, 2003. xiv, 300 p. Cloth $49.95

DEELY, JOHN: *The Impact on Philosophy of Semiotics.* South Bend, IN: St Augustine's, 2003. viii, 267 p. Cloth $26.00.

DEWITT, RICHARD: *Worldviews: An Introduction to the History and Philosophy of Science.* Malden, MA: Blackwell, 2004. xii, 326 p. Cloth $64.95, paper $29.95.

DICKERSON, A.B.: *Kant on Representation and Objectivity.* New York: Cambridge, 2004. x, 217 p. Cloth $60.00.

DILLON, JOHN, and LLOYD P. GERSON: *Neoplatonic Philosophy: Introductory Readings.* Indianapolis: Hackett, 2004. xxiv, 373 p. Cloth $47.95, paper $18.95.

DOMBROWSKI, DANIEL A.: *Divine Beauty: The Aesthetics of Charles Hartshorne.* Nashville: Vanderbilt, 2004. Cloth $59.95.

DROEGE, PAULA: *Caging the Beast: A Theory of Sensory Consciousness.* Philadelphia: John Benjamins, 2003. x, 183 p. Paper $54.95.

DUMMETT, MICHAEL: *Truth and the Past.* New York: Columbia, 2004. x, 122 p. Cloth $29.50.

EDMUNDSON, WILLIAM A.: *An Introduction to Rights.* New York: Cambridge, 2004. xv, 223 p. Cloth $55.00, paper $22.00.

ELDER, CRAWFORD L.: *Real Nature and Familiar Objects.* Cambridge: MIT, 2004. Cloth $32.00.

ELDRIDGE, RICHARD: *An Introduction to the Philosophy of Art.* New York: Cambridge, 2003. ix, 285 p. Cloth $65.00, paper $23.00.

ENÇ, BERENT: *How We Act: Causes, Reasons, and Intentions.* New York: Oxford, 2003. xiii, 252 p. Cloth $45.00.

FALK, ARTHUR E.: *Desire and Belief: Introduction to Some Recent Philosophical Debates.* Lanham, MD: UP of America, 2004. xxi, 595 p. Paper $62.00.

FINE, GAIL: *Plato on Knowledge and Forms: Selected Essays.* New York: Oxford, 2003. ix, 447 p. Cloth $99.00.

FINKELSTEIN, DAVID H.: *Expression and the Inner.* Cambridge: Harvard, 2003. ix, 182 p. Cloth $45.00.

FLEISCHACKER, SAMUEL: *On Adam Smith's Wealth of Nations*: *A Philosophical Companion.* Princeton: University Press, 2004. xvii, 329 p. Cloth $39.50.

FLEMING, BRUCE E.: *Art and Argument: What Words Can't Do and What They Can.* Lanham, MD: UP of America, 2003. vi, 201 p. Paper $34.00.

FODOR, JERRY A.: *Hume Variations.* New York: Oxford, 2003. 165 p. Cloth $22.00.

FOSTER, JOHN: *The Divine Lawmaker: Lectures on Induction, Laws of Nature, and the Existence of God.* New York: Oxford, 2004. viii, 191 p. Cloth $45.00.

FRANKFURT, HARRY G.: *The Reasons of Love.* Princeton: University Press, 2004. 100 p. Cloth $19.95.

GANDER, ERIC M.: *On Our Minds: How Evolutionary Psychology Is Reshaping the Nature-versus-Nurture Debate.* Baltimore: Johns Hopkins, 2003. x, 293 p. Cloth $45.00.

GÄRDENFORS, PETER: *How Homo Became Sapiens: On the Evolution of Thinking.* New York: Oxford, 2003. ix, 240 p. Cloth $59.50.

GARVER, EUGENE: *For the Sake of Argument: Practical Reasoning, Character, and the Ethics of Belief.* Chicago: University Press, 2004. xi, 272 p. Paper $22.50.

GIER, NICHOLAS F.: *The Virtue of Nonviolence: From Gautama to Gandhi.* Albany: SUNY, 2004. xv, 224 p. Cloth $40.00.

GOLDFARB, WARREN: *Deductive Logic.* Indianapolis: Hackett, 2003. xv, 292 p. Cloth $32.50.

GOLOMB, JACOB: *Nietzsche and Zion.* Ithaca, NY: Cornell, 2004. xii, 274 p. Cloth $39.95.

GRANT, ROBERT: *Imagining the Real: Essays on Politics, Ideology, and Literature.* New York: Palgrave MacMillan, 2003. xvi, 248 p. Cloth $69.95.

GREEN, CELIA: *The Lost Cause: Causation and the Mind-Body Problem.* Cambridge, UK: Oxford Forum, 2003. xli, 241 p. Cloth.

GREEN, T.H.: *Prolegomena to Ethics.* DAVID O. BRINK, ed. New York: Oxford, 2003. cl, 477 p. Cloth $85.00, paper $29.95.

GIULIANI, MASSIMO: *A Centaur in Auschwitz: Reflections on Primo Levi's Thinking.* Lanham, MD: Lexington, 2003. ix, 111 p. Cloth $60.00, paper $24.95.

HALPERN, JOSEPH Y.: *Reasoning about Uncertainty.* Cambridge: MIT, 2003. xiv, 483 p. Cloth $50.00.

HALWANI, RAJA: *Virtuous Liaisons: Care, Love, Sex, and Virtue Ethics.* Chicago: Open Court, 2003. ix, 315 p. Paper $27.95.

HAMBLET, WENDY C.: *The Sacred Monstrous: A Reflection on Violence in Human Communities.* Lanham, MD: Lexington, 2004. xiv, 111 p. Cloth $65.00, paper $22.00.

HANNA, PATRICIA, and BERNARD HARRISON: *Word and World: Practice and the Foundations of Language.* New York: Cambridge, 2004. xii, 420 p. Cloth $85.00, paper $30.00.

HARRIS, LEE: *Civilization and Its Enemies: The Next Stage of History.* New York: Free Press, 2004. xix, 232 p. Cloth $26.00.

HASKINS, EKATERINA V.: *Logos and Power in Isocrates and Aristotle.* Columbia: South Carolina UP, 2004. xiii, 172 p. Cloth $34.95.

HAWTHORNE, JOHN: *Knowledge and Lotteries.* New York: Oxford, 2004. viii, 205 p. Cloth $35.00.

HERWITZ, DANIEL: *Race and Reconciliation.* Minneapolis: Minnesota UP, 2003. xvii, 227 p. Cloth $60.95, paper $21.95.

HETCHER, STEVEN A.: *Norms in a Wired World.* New York: Cambridge, 2004. xii, 420 p. Cloth $65.00.

HILLS, HELEN: *Invisible City: The Architecture of Devotion in Seventeenth-Century Neapolitan Convents.* New York: Oxford, 2004. xiii, 268 p. Cloth $60.00.

HIMMELFARB, GERTRUDE: *The New History and the Old: Critical Essays and Reappraisals,* Revised Edition. Cambridge: Harvard, 2004. 266 p. Paper $18.95.

HOY, DAVID COUZENS: *Critical Resistance: From Poststructuralism to Post-Critique.* Cambridge: MIT, 2004. viii, 274 p. Cloth $35.00.

HUGHES, CHRISTOPHER: *Kripke: Names, Necessity, and Identity.* New York: Oxford, 2004. xii, 247 p. Cloth $60.00.

JACKSON, FRANK, and PHILIP PETTIT, and MICHAEL SMITH: *Mind, Morality, and Explanation: Selected Collaborations.* New York: Oxford, 2004. xii, 427 p. Cloth $95.00.

JACOB, PIERRE, and MARC JEANNEROD: *Ways of Seeing: The Scope and Limits of Visual Cognition.* New York: Oxford, 2003. xix, 287 p. Cloth $124.50.

JACOBS, LESLEY A.: *Pursuing Equal Opportunities: The Theory and Practice of Egalitarian Justice.* New York: Cambridge, 2003. Cloth $65.00, paper $24.00.

JACQUETTE, DALE: *Pathways in Philosophy: An Introduction Guide with Readings.* New York: Oxford, 2003. xv, 555 p. Paper $56.95.

JAMES, WENDY: *The Ceremonial Animal: A New Portrait of Anthropology.* New York: Oxford, 2003. xxiii, 384 p. Cloth $45.00.

JANOWSKI, ZBIGNIEW: *Augustinian-Cartesian Index: Texts and Commentary.* South Bend, IN: St. Augustine's, 2004. xv, 275 p. Cloth $35.00.

JOHNSON, DAVID MARTEL: *How History Made the Mind: The Cultural Origins of Objective Thinking.* Chicago: Open Court, 2003. xi, 227 p. Paper $24.95.
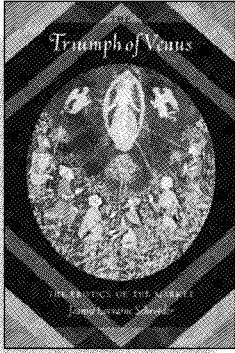
KATZ, CLAIRE ELISE: *Levinas, Judaism, and the Feminine: The Silent Footsteps of Rebecca.* Bloomington: Indiana UP, 2003. xvi, 199 p. Cloth $49.95, paper $22.95.

KATZ, JERROLD J.: *Sense, Reference, and Philosophy.* New York: Oxford, 2004. xvi, 224 p. Cloth $49.95.

KEKES, JOHN: *The Illusions of Egalitarianism.* Ithaca: Cornell, 2003. xii, 228 p. Cloth $29.95.

KELLY, EUGENE: *The Basics of Western Philosophy.* Westport, CT: Greenwood, 2004. xii, 295 p. Cloth $75.00.

KOLNAI, AUREL: *On Disgust.* BARRY SMITH and CAROLYN KORSMEYER, eds. Chicago, Open Court, 2004. viii, 120 p. Paper $18.95.

KRAMER, MATTHEW H.: *The Quality of Freedom.* New York: Oxford, 2003. ix, 482 p. Cloth $45.00.

KUKATHAS, CHANDRAN: *The Liberal Archipelago: A Theory of Diversity and Freedom.* New York: Oxford, 2003. xii, 292 p. Cloth $45.00.

KÜNNE, WOLFGANG: *Conceptions of Truth.* New York: Oxford, 2003. xiii, 493 p. Cloth $74.00.

KUSPIT, DONALD: *The End of Art.* New York: Cambridge, 2004. xi, 208 p. Cloth $28.00.

LANDMAN, FRED: *Indefinites and the Type of Sets.* Malden, MA: Blackwell, 2004. xxvii, 269 p. Cloth $78.95, paper $43.95.

LAPORTE, JOSEPH: *Natural Kinds and Conceptual Change.* New York: Cambridge, 2004. x, 221 p. Cloth $70.00.

LAWLOR, LEONARD: *The Challenge of Bergsonism: Phenomenology, Ontology, Ethics.* New York: Continuum, 2003. xiii, 154 p. Cloth $105.00, paper $24.95.

LEAR, GABRIEL RICHARDSON: *Happy Lives and the Highest Good: An Essay on Aristotle's Nichomachean Ethics.* Princeton: University Press, 2004. ix, 238 p. Cloth $35.00.

LEWENS, TIM: *Organisms and Artifacts: Design in Nature and Elsewhere.* New York: Cambridge, 2004. xi, 183 p. Cloth $32.00.

LIU, SHU-HSIEN: *Essentials of Neo-Confucian Philosophy.* Westport, CT: Greenwood, 2003. xv, 167 p. Cloth $89.95.

LLEWELYN, JOHN: *Seeing through God: A Geophenomenology.* Bloomington: Indiana UP, 2004. xv, 193 p. Cloth $59.95, paper $24.95.

LLOYD, DAN: *Radiant Cool: A Novel Theory of Consciousness.* Cambridge: MIT, 2004. xvii, 357 p. Cloth $24.95.

LLOYD, G.E.R.: *Ancient Worlds, Modern Reflections: Philosophical Perspectives on Greek and Chinese Science and Culture.* New York: Oxford, 2004. xi, 222 p. Cloth $35.00.

LUCY, NIALL: *A Derrida Dictionary.* Malden, MA: Blackwell, 2004. xii, 183 p. Paper $24.95.

LUNTLEY, MICHAEL: *Wittgenstein: Meaning and Judgement.* Malden, MA: Blackwell, 2003. viii, 187 p. Paper $29.95.

MCFEE, GRAHAM: *Sport, Rules, and Values: Philosophical Investigations into the Nature of Sport.* New York: Routledge, 2004. x, 2002 p. Paper $24.99.

MCINERNY, RALPH: *Aquinas.* New York: Polity, 2004. vii, 160 p. Cloth $59.95,

paper $24.95; *The Very Rich Hours of Jacques Maritain: A Spiritual Life.* Notre Dame: University Press, 2003. viii, 235 p. Cloth $32.00.

MELNYK, ANDREW: *A Physicalist Manifesto: Thoroughly Modern Materialism.* New York: Cambridge, 2003. xii, 328 p. Cloth $70.00.

MEYERS, DIANA TIETJENS: *Essays on Identity, Action, and Social Life.* Lanham, MD: Rowman and Littlefield, 2004. xix, 319 p. Cloth $75.00, paper $27.95.

MOLNAR, GEORGE: *Powers: A Study in Metaphysics.* STEPHEN MUMFORD, ed. New York: Oxford, 2003. xiv, 238 p. Cloth $45.00.

MONTUSCHI, ELEONORA: *The Objects of Social Science.* New York: Continuum, 2003. ix, 160 p. Paper $29.95.

MORIARTY, MICHAEL: *Early Modern French Thought: The Age of Suspicion.* New York: Oxford, 2003. xii, 271 p. Cloth $75.00.

MORTON, ADAM: *Philosophy in Practice: An Introduction to the Main Questions,* Second Edition. Malden, MA: Blackwell, 2004. xv, 446 p. Paper $39.95.

NANCY, JEAN-LUC: *A Finite Thinking.* SIMON SPARKS, ed. Stanford: University Press, 2003. viii, 351 p. Cloth $60.00, paper $24.95.

NICHOLS, SHAUN, and STEPHEN P. STICH: *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds.* New York: Oxford, 2003. 237 p. Cloth $70.00, paper $24.95.

NOONAN, HARRY W.: *Personal Identity.* New York: Routledge, 2003. xiii, 239 p. Paper $31.95.

NUSSBAUM, MARTHA C.: *Hiding from Humanity: Disgust, Shame, and the Law.* Princeton: University Press, 2004. xv, 413 p. Cloth $29.95.

OLIN, DORIS: *Paradox.* Ithaca: McGill-Queen's, 2003. x, 222 p. Cloth $70.00, paper $27.95.

O'MEARA, DOMINIC J.: *Platonopolis: Platonic Political Philosophy in Late Antiquity.* New York: Oxford, 2003. xi, 249 p. Cloth $55.00.

OTT, WALTER R.: *Locke's Philosophy of Language.* New York: Cambridge, 2004. vii, 159 p. Cloth $55.00.

PARKER, FRED: *Scepticism and Literature: An Essay on Pope, Hume, Sterne, and Johnson.* New York: Oxford, 2003. x, 290 p. Cloth $72.00.

PARTEE, BARBARA H.: *Compositionality in Formal Semantics: Selected Papers by Barbara H. Partee.* Malden, MA: Blackwell, 2004. xi, 331 p. Cloth $79.95, paper $44.95.

PASKOW, ALAN: *The Paradoxes of Art: A Phenomenological Investigation.* New York: Cambridge, 2004. xi, 260 p. Cloth $65.00.

PEACOCKE, CHRISTOPHER: *The Realm of Reason.* New York: Oxford, 2004. x, 284 p. Cloth $35.00.

PIHLSTÖM, SAMI: *Naturalizing the Transcendental: A Pragmatic View.* Amherst, NY: Humanity, 2003. 390 p. Cloth $49.00.

PLACE, U.T.: *Identifying the Mind: Selected Papers of U.T. Place.* GEORGE GRAHAM, and ELIZABETH R. VALENTINE, eds. New York: Oxford, 2004. xii, 215 p. Cloth $35.00.

POLGER, THOMAS W.: *Natural Minds.* Cambridge: MIT, 2004. xxvii, 294 p. Cloth $38.00.

POTTER, MICHAEL: *Set Theory and Its Philosophy: A Critical Introduction.* New York: Oxford, 2004. xiii, 345 p. Cloth $85.00, paper $29.95.

PRIOR, ARTHUR N.: *Papers on Time and Tense.* PER HASLE, PETER ØHRSTRØM, TORBEN BRAÜNER, and JACK COPELAND, eds. New York: Oxford, 2003. ix, 331 p. Cloth $72.00, paper $24.95.

PUTNAM, HILARY: *Ethics without Ontology.* Cambridge: Harvard, 2004. x, 160 p. Cloth $29.95.

PYLYSHYN, ZENON: *Seeing and Visualizing: It's Not What You Think.* Cambridge: MIT, 2003. xviii, 563 p. Cloth $50.00.

QUINE, W.V.: *Quintessence: Basic Readings from the Philosophy of W.V. Quine.* ROGER F. GIBSON, JR., ed. Cambridge: Harvard, 2004. x, 416 p. Cloth $39.95.

RAATZSCH, RICHARD: *Eigentlich Seltsames: Wittgensteins "Philosophiche Untersuchungen": Band I: Einleitung und Kommentar PU 1-64.* Paderborn, Germany: Schöningh, 2003. 490 p. Cloth.

REGAN, TOM: *Empty Cages: Facing the Challenge of Animal Rights.* Lanham, MD: Rowman and Littlefield, 2004. xi, 229 p. Cloth $21.95.

ROBERT, JASON SCOTT: *Embryology, Epigenesis, and Evolution: Taking Development Seriously.* New York: Cambridge, 2004. xvi, 158 p. Cloth $60.00.

ROTHSTEIN, SUSAN: *Structuring Events: A Study in the Semantics of Lexical Aspect.* Malden, MA: Blackwell, 2004. x, 206 p. Paper $43.95.

RUNDLE, BEDE: *Why There Is Something Rather than Nothing.* New York: Oxford, 2004. xii, 204 p. Cloth $45.00.

SAID, EDWARD W.: *Humanism and Democratic Criticism.* New York: Columbia, 2004. xvii, 154 p. Cloth $19.95.

SAMUELSON, NORBERT M.: *Jewish Philosophy: An Historical Introduction.* New York: Continuum, 2003. vii, 344 p. Cloth $125.00, paper $34.95.

SÁNCHEZ, GONZALO J., JR.: *Pity in Fin-de-Siècle French Culture: "Liberté, Egalité, Pitié."* Westport, CT: Praeger, 2004. x, 319 p. Cloth $67.95.

SANTAYANA, GEORGE: *The Letters of George Santayana: Book Five, 1933–1936.* WILLIAM G. HOLZBERGER, ed. Cambridge: MIT, 2003. lxxvii, 609 p. Cloth $60.00.

SCHAUER, FREDERICK: *Profiles, Probabilities, and Stereotypes.* Cambridge: Harvard, 2003. xiii, 359 p. Cloth $29.95.

SCHIFFER, STEPHEN: *The Things We Mean.* New York: Oxford, 2003. 362 p. Cloth $74.00, paper $24.95.

SHAFER-LANDAU, RUSS: *Moral Realism: A Defense.* New York: Oxford, 2003. x, 322 p. Cloth $45.00.

SHAMDASANI, SONU: *Jung and the Making of Modern Psychology: The Dream of a Science.* New York: Cambridge, 2003. xvi, 387 p. Cloth $75.00, paper $28.00.

SHANKS, NIALL: *God, the Devil, and Darwin: A Critique of Intelligent Design Theory.* New York: Oxford, 2004. xiii, 273 p. Cloth $29.95.

SHAPIRO, LAWRENCE A.: *The Mind Incarnate.* Cambridge: MIT, 2004. xvi, 253 p. Cloth $34.00.

SHOEMAKER, SYDNEY: *Identity, Cause, and Mind*, Expanded Edition. New York: Oxford, 2003. x, 461 p. Paper $24.95.

SINGER, IRVING: *Three Philosophical Filmmakers: Hitchcock, Welles, Renoir.* Cambridge: MIT, 2004. x, 279 p. Cloth $32.95.

SOAMES, SCOTT: *Philosophical Analysis in the Twentieth Century,* Volume 2: *The Age of Meaning.* Princeton: University Press, 2003. xxii, 479 p. Cloth $35.00.

SOBEL, JORDAN HOWARD: *Logic and Theism: Arguments for and against Beliefs in God.* New York: Cambridge, 2004. xix, 652 p. Cloth $95.00.

SOLOMON, ROBERT C: *Living with Nietzsche: What the Great "Immoralist" Has to Teach Us.* New York: Oxford, 2003. x, 243 p. Cloth $35.00.

SOLSO, ROBERT L.: *The Psychology of Art and the Evolution of the Conscious Brain.* Cambridge, MIT, 2003. xvi, 278 p. Cloth $45.00.

SORENSEN, ROY: *A Brief History of the Paradox: Philosophy and the Labyrinths of the Mind.* New York: Oxford, 2003. xv, 394 p. Cloth $26.00.

SPITZER, MICHAEL: *Metaphor and Musical Thought.* Chicago: University Press, 2004. x, 380 p. Cloth $60.00.

STALEY, KENT W.: *The Evidence for the Top Quark: Objectivity and Bias in Collaborative Experimentation.* New York: Cambridge, 2004. xvi, 343 p. Cloth $70.00.

STALNAKER, ROBERT C.: *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays.* New York: Oxford, 2003. xii, 287 p. Cloth $72.00, paper $22.00.

STEWART, JON: *Kierkegaard's Relations to Hegel Reconsidered.* New York: Cambridge, 2003. xix, 695 p. Cloth $55.00.

STOUT, JEFFREY: *Democracy and Tradition.* Princeton: University Press, 2004. xv, 349 p. Cloth $35.00.

SULLIVAN, VICKIE B.: *Machiavelli, Hobbes, and the Formation of a Liberal Republicanism in England.* New York: Cambridge, 2004. x, 284 p. Cloth $75.00.

TALLIS, RAYMOND: *The Hand: A Philosophical Inquiry into Human Being.* Edinburgh: University Press, 2003. x, 364 p. Cloth $90.00, paper $30.00.

TAN, SOR-HOON: *Confucian Democracy: A Deweyan Reconstruction.* Albany: SUNY, 2003. viii, 258 p. Cloth $65.50, paper $20.95.

TAYLOR, CHARLES: *Modern Social Imaginaries.* Durham, NC: Duke, 2004. xi, 215 p. Paper $18.95.

TAYLOR, JOAN E.: *Jewish Women Philosophers of First-Century Alexandria: Philo's "Therapeutae" Reconsidered.* New York: Oxford, 2003. xv, 417 p. Cloth $99.00.

TAYLOR, PAUL C.: *Race: A Philosophical Introduction.* Malden, MA: Polity, 2004. xiv, 202 p. Cloth $62.95, paper $26.95.

TYE, MICHAEL: *Consciousness and Persons: Unity and Identity.* Cambridge: MIT, 2003. xv, 203 p. Cloth $35.00.

WARD, ROGER A.: *Conversion in American Philosophy: Exploring the Practice of Transformation.* New York: Fordham, 2004. xli, 235 p. Cloth $70.00.

WERTHEIMER, ALAN: *Consent to Sexual Relations.* xv, 293 p. New York: Cambridge, 2003. Cloth $70.00, paper $26.00.

WEST, HENRY R.: *An Introduction to Mill's Utilitarian Ethics.* New York: Cambridge, 2004. xi, 216 p. Cloth $55.00, paper $20.00.

WETTSTEIN, HOWARD: *The Magic Prism: An Essay in the Philosophy of Language.* New York: Oxford, 2004. x, 241 p. Cloth $49.95.

THE JOURNAL OF PHILOSOPHY

The Journal is pleased to announce to our readers and subscribers the launch of its new website, which will have the pages listed below. The address is: *www.journalofphilosophy.org* or simply *www.jphil.org*.

∼ • ∼

PURPOSE & HISTORY
CONTACTS AND ADDRESSES
BOARD OF TRUSTEES
EDITORIAL BOARD
MEMOS TO AUTHORS
SUBSCRIPTIONS & ADVERTISING

RECENT & FORTHCOMING ISSUES
ELECTRONIC ISSUES
INDEX 1904–PRESENT
RIGHTS AND PERMISSIONS
LECTURES & CONFERENCES
PHILOSOPHY LINKS

∼ • ∼