

5: Preliminaries to the Argument

In this chapter, we set forth the logical structure of the argument we will use in chapter six in our attempt to show that *Nfc* is self-refuting. Thus, our main topics in this chapter are self-referential statements and the arguments which are based upon the ways in which these statements go wrong.

As necessary background for this discussion of self-reference, we devote the first section of this chapter to the common ways in which statements can go wrong. We also consider at the end of the chapter that aspect of the *PNfc*'s position which gives rise to the self-referential difficulties we will show in chapter six.

Thus, this chapter has six sections: A) How statements go wrong; B) Self-referential statements; C) How self-referential statements go wrong; D) The falsification of self-referential statements; E) The affirmation of "No free choice"; and F) Rationality norms as conditions for affirmation.

A. How statements go wrong

Since philosophical theories are statements, one can find out how the former go wrong by considering how the latter go wrong. We define a statement—in the sense of a *stating*—as follows: A statement (*S*) includes someone's act (*A*) of affirming a proposition (*P*) by way of a token (usually a sentence) (*T*). A statement, then, includes an affirmation—in the sense of an act of affirming. By "affirmation" we mean all those acts in which someone claims—with whatever degree of conviction—that a proposition is true. Thus, hypotheses and opinions frequently are affirmed. For example, to claim that one position is more reasonable than another is to affirm that position. Commands, questions, prayers, and fictional pretenses are not affirmations. We call affirming, denying, questioning, and other acts upon propositions "propositional acts."¹

We also distinguish between the act of making a statement, what is stated, and the sentence used to express what is stated. We use the expression “performance of the statement” to refer to the act of affirming what is stated and to the act of uttering the sentence. We use the word “proposition” to refer to what is stated. A proposition has both sense and reference; it is either true or false. We use the word “sentence” to refer to the linguistic entity the utterance of which—by speech or by writing or otherwise—expresses the proposition which is affirmed.

Statements can go wrong because of failure of the proposition, or of the sentence, or of the performance of the statement.

Propositions can be false. The state of affairs picked out by a proposition might not obtain.

Propositions considered in sets can go wrong in another way. Sets of propositions can be formally inconsistent—that is, they can be reduced to propositions of the form p and $not-p$. Since philosophical theories are made up of a number of propositions—some of them unstated assumptions, some of them remote implications—the ordinary tests of formal logic are very important in philosophical discussion. It is worth noting that the discovery of a formal inconsistency does not require one to abandon any *particular* proposition. Since inconsistency obtains between two propositions, one can remove the inconsistency by surrendering either of them.

Statements also can go wrong by a failure of their sentences. A sentence can fail to express a proposition. Frequently, sentences of this sort are called “meaningless.” We use “meaningless” to mean a defect either of sense or of reference.

One common way in which a sentence can fail to be meaningful is by being vague or indefinite. Vagueness and indefiniteness are semantic difficulties, because they prevent determinate reference.

Semantic difficulties such as vagueness and indefiniteness are often pointed out by critics of philosophical theories. Perhaps the most common use of the philosopher’s question, “What exactly do you mean?” is to demand that the state of affairs under consideration be delineated precisely and fully. Our first chapter exemplifies such semantic clarification. *Sfc/Nfc* has often been confused because of the different meanings of “free.” To avoid this confusion these meanings are distinguished. The controversy also has been impeded by the indefiniteness of the expression “free choice.” To remedy this indefiniteness the empirical reference of the notion is clarified.

There are philosophical arguments which claim to show that certain philosophical theses have *inherent* semantic difficulties. If such an argument is successful, it warrants the charge that the thesis in question is meaningless. We used such an argument in chapter four in our criticism of the double-aspect theory.

Both Plato's and Aristotle's arguments against the theory that all is flux, and Aristotle's argument against the theory that all is undifferentiated unity are attempts to show that these theories render reference impossible. In a world of undifferentiated unity or in a world of complete flux, no state of affairs could obtain. It would be impossible to refer to such a world.²

The examples drawn from Plato and Aristotle are of arguments which seek to show that the theories criticized entail the impossibility of reference—including the reference of these very theories. Our argument against the double-aspect theory showed that this theory either leads to formal inconsistency or makes it impossible to refer to some of the phenomena for which the theory is intended to account.

Statements also can go wrong by a failure of their performances. A performance of a statement goes wrong when the performance cannot achieve its purpose—that is, when the performance is pointless. A performance can be pointless for a number of different reasons. One of these is if the statement lacks a suitable context. As C. K. Grant indicates, utterances must take place in a context—that is, they imply certain propositions about their speaker, their audience, and so on. If the implied propositions are false, then the utterance is irrational. According to Grant, the performance of making a statement has psychological and propositional “pragmatic implications.” For example, one using an indicative sentence ordinarily implies that he believes it, that he has an audience, and that he wishes the audience to believe it. These implications do not follow from the proposition affirmed; yet if they are false, the utterance is *pointless*.³

Performative pointlessness not only is common in the ordinary usage of speech acts of various kinds; it also explains a philosophically important informal fallacy—that of begging the question. There is nothing formally wrong with the circularity involved in this fallacy. The fallacy arises because the circularity makes impossible the successful achievement of the purpose of the argument. The purpose of an argument is to make its conclusions rationally acceptable by virtue of its relation to grounds which are independently acceptable. If the truth of the conclusion is assumed as part of the premises, then the argument cannot achieve its desired effect of *making* the conclusion acceptable on the basis of something independently acceptable. We have shown in chapters two and three that many arguments for and against *Sfc* are pointless in just this way.

B. Self-referential statements

Self-referential statements can go wrong in all the ways other statements can go wrong. Their self-referential character, however, makes a difference to the logical properties of their going wrong and this difference is philosophically

important. Because the notion of self-reference is tricky, we define it and give examples of its various kinds. We also argue that there is nothing inherently wrong with self-referential statements.

A *statement* is self-referential if and only if the proposition which is affirmed refers to some aspect of the statement—that is, either to the sentence, or to the performance of affirming or uttering, or to the proposition itself.

We first consider self-referential statements in which the reference is to the sentence. Smith's statement (*S*), "All Smith's English statements are cases of correct English," affirms a general proposition. Each instance of this general proposition is a proposition which, 1) refers to the sentence used to make a statement of Smith's in English, and, 2) says of it that it is a case of correct English. Let "*T*" name the sentence which expresses the proposition affirmed in *S*. This proposition has a self-referential instance: "*T* is the sentence used to make a statement of Smith's in English and *T* is a case of correct English." We call the self-reference of this statement "sentential." The proposition in *S* would not be sententially self-referential if Smith used a French sentence to make his statement. Nor would the proposition in *S* be sententially self-referential if someone else said: "All Smith's English statements are cases of correct English."

In a similar way, Smith's statement, "All Smith's statements are cases of correct English," is sententially self-referential. The proposition affirmed in this statement is sententially self-referential whenever it is affirmed by Smith. Likewise, "All my statements are cases of correct English" is a sententially self-referential statement. The propositions affirmed in statements made by using these words will be different for each speaker, since "my" will have a different referent in each. Yet the propositions affirmed by these words will all be sententially self-referential, whenever a speaker uses these words to make a statement.

There are self-referential statements in which the reference is to the performance. We call these "performatively self-referential statements." For example, Smith's spoken statement (*S*), "Smith always speaks softly," is performatively self-referential. Let "*U*" name Smith's speaking the sentence used to express the proposition affirmed in *S*. This proposition has a self-referential instance: "*U* is a spoken utterance of Smith's and *U* is spoken softly." The proposition affirmed in *S* need not be performatively self-referential; Smith might express in writing the proposition expressed in *S*, or someone else might express this proposition.

There also are performatively self-referential statements in which the reference is to the affirming rather than to the uttering. For example, Smith's statement, "All Smith's statements are well-founded," is performatively self-referential. The reference is to Smith's affirming.

There also are self-referential statements in which the reference is to the

proposition. We call these “semantically self-referential statements.” For example, Smith’s statement (*S*), “All Smith’s statements are contingently true,” is semantically self-referential. Let “*P*” name the general proposition affirmed in *S*. Each instance of this general proposition is a proposition, 1) referring to some proposition affirmed by Smith, and, 2) saying of it that it is contingently true. The proposition affirmed in *S* has a semantically self-referential instance, “*P* is a proposition and *P* is contingently true.” This statement is semantically self-referential only when affirmed by Smith.

The previous examples might suggest that all self-referential statements contain general propositions. This need not be so. For example, “This sentence contains five words” can be used to express a sententially self-referential statement. “This sentence is printed on paper” can be used to express a performatively self-referential statement. However, semantically self-referential statements cannot be singular; sentences which seem to express the propositions in singular semantically self-referential statements turn out to be paradoxical. We discuss these below.

Thus far we have been considering *statements* which are self-referential. As the examples make clear, the propositions affirmed in these statements are not inherently self-referential—that is, the propositions might be expressed in ways which would not involve self-reference. For instance, the propositions might be affirmed by different speakers, or by way of different sentences.

However, there are certain propositions which are self-referential in any possible stating of them. Such propositions inevitably refer to themselves or to the performances of their statements or to the sentences by which they are expressed. We call such propositions “self-referential propositions”; they should not be confused with the propositions of self-referential statements.

The following are examples of self-referential propositions. “All statements can be made in English” expresses a sententially self-referential proposition. “All affirmations are well-founded” expresses a performatively self-referential proposition. “All propositions are either true or false” expresses a semantically self-referential proposition.

In each of these examples an instance of the general proposition refers to the sentence or performance or proposition. No matter how these propositions are stated, they make reference to the sentence, performance, or proposition. In other words, the self-reference of these propositions is invariant in relation to any change of sentences used to express them, persons who affirm them, and modalities of uttering and affirming them.

As already noted, semantically self-referential statements cannot be singular. Statements such as “This very statement is vague” and “This very statement is false” are puzzling. One wonders just what proposition it is of which the vagueness or falsity is predicated. In fact, “This very statement is false” is a version of the Liar Paradox. If one takes it to be true, then it turns out false,

since it says that it is false. If, on the other hand, one assumes it to be false, then it turns out true. Other self-referential statements also lead to paradox—for example, Grelling's paradox concerning "heterologicality." These have been called "semantic paradoxes." Seemingly similar paradoxes arise in the use of certain formal notions to refer to themselves. These are called "logical paradoxes." Russell's paradox concerning the class of all classes which are not members of themselves is a famous example.

The semantic and logical paradoxes of self-reference have given rise to the claim that all self-reference is illegitimate and leads to paradox. We dispute this claim. Analysis of self-referring statements which do lead to paradox does not support the claim that all self-reference is illegitimate; rather the kinds of self-reference which are illegitimate can be shown to be so for specifiable reasons.

The claim that all self-reference must be avoided was first articulated by Russell and Whitehead in *Principia Mathematica*. They contend that all self-referential statements are nonsensical since they all violate what Russell and Whitehead call "the vicious circle principle."

All self-referential statements, they argue, violate this principle by making reference to "illegitimate totalities," such as the one referred to in the paradox of the class of all classes which are not members of themselves. Thus, "All propositions are either true or false" is rejected as nonsensical, as is the statement of the Liar Paradox.⁴ Likewise, they apply their principle to the well-known argument that skepticism is self-refuting.

Similarly, the imaginary sceptic, who asserts that he knows nothing, and is refuted by being asked if he knows that he knows nothing, has asserted nonsense, and has been fallaciously refuted by an argument which involves a vicious-circle fallacy. In order that the sceptic's assertion may become significant, it is necessary to place some limitation upon the things of which he is asserting his ignorance, because the things of which it is possible to be ignorant form an illegitimate totality. But as soon as a suitable limitation has been placed by him upon the collection of propositions of which he is asserting his ignorance, the proposition that he is ignorant of every member of this collection must not itself be one of the collection. Hence any significant scepticism is not open to the above form of refutation.⁵

Thus, *all* self-reference is rejected as illegitimate.

This view has been further articulated by those who regard language as a hierarchy of levels each of which can have as referents only linguistic and nonlinguistic entities lower in the hierarchy. Moreover, the view that all self-reference is illegitimate has been taken as a basis for regarding all arguments which depend on self-reference—including those against *Nfc*—as fallacious.⁶

This claim that all self-reference is illegitimate and must be avoided involves

a number of widely recognized difficulties.⁷ Several of these difficulties are crucial.

The prohibition of self-reference may be stated as follows: "There are no self-referential propositions." This formulation is logically equivalent to (*S*): "All propositions are non-self-referential." If the utterance of *S* expresses the affirmation of a proposition (*P*), then this proposition is a general one, whose instances taken collectively make reference to all propositions. One of these instances makes reference to *P*; this instance is "*P* is a proposition and *P* is non-self-referential." In other words, *P* is semantically self-referential in virtue of the fact that it is about *all* propositions. However what *P* says about all propositions, including itself, is that they are non-self-referential. It follows that either *S* does not express a proposition, or *P* falsifies itself.

On the one hand, if *S* does not express a proposition, then *S* can make no truth-claim; it cannot *deny* that there are propositions which contain self-reference. Thus *S* has no cognitive purpose; it is simply not a truth-claim. On the other hand, if a proposition is affirmed in *S*, then the proposition is self-referential and falsifies itself. Any attempt to avoid this dilemma by limiting the scope of the prohibition would allow for *some* self-reference.⁸

The second decisive objection to the complete prohibition of self-reference as illegitimate is that the prohibition ignores the differences between kinds of self-reference. Russell and Whitehead regard as identical in kind the self-reference of the application of formal notions to themselves, the self-reference of the semantical paradoxes, and the performative self-reference of skepticism. These are clearly different, as are the paradoxes which arise in each case.

Logical paradoxes, such as Russell's, are not genuine antinomies by which one is driven to hold both elements of a contradiction. The class of all classes which are not members of themselves will be a member of itself if one assumes that it is not; it will not be a member of itself if one assumes that it is. But one can remove the paradox by noticing that it depends upon the assumption that there is a class of all classes which are not members of themselves. If, rather than making this assumption, one asks whether there could be such a class, what was stated as a paradox can be seen to be a *reductio ad absurdum* argument to support a negative answer to the question. This argument shows that the notion of a class of all classes which are not members of themselves is incoherent. There can be no such class.⁹ There appears to be a paradox only because of the way the issue is formulated. That there is such a class is erroneously taken as given, not stated as an assumption to be proved or disproved.

Quine points out that semantic paradoxes, unlike logical paradoxes, do lead to genuine antinomies.¹⁰ These too, however, depend upon an assumption, although one of a different sort than in the logical paradoxes. The assumption

on which all the semantic paradoxes rest is that the sentences which generate them express propositions. This assumption is false.

With regard to the putative statement, "This very statement is false," one can ask what proposition it is of which the falsity is predicated. A similar question can be asked with regard to such puzzling sentences as "This very statement is true" and "This very statement is vague." Ordinarily, when someone asks for semantic clarification, a straightforward answer is possible; the speaker can indicate what he is talking about. The speaker can name it or pick it out by a definite description. This is also the case when the object referred to is a proposition. Thus, if someone asks which proposition a speaker is claiming to be true, or vague, or false; the speaker can express the proposition in some other way, name it, or describe it.

These procedures fail for statements which lead to semantic paradoxes. If someone says "This very statement is false" and is asked what proposition he is talking about, his answer, "This very statement," or his repeating the statement, or anything else he might do is no help since such a reply cannot indicate a referent. Unless it can be shown that the subject, "This very statement," does in fact refer to a proposition, then this sentence does not express a proposition—it has no reference.

For example, someone might ask whether "This very statement" in "This very statement is false" has a referent. Any attempt to show that the referring expression, "This very statement," does in fact refer to a proposition expressed in "This very statement is false" gives rise to the same question. Thus, "This very statement" might be said to refer to "This very statement is false" giving rise to "The statement, 'This very statement is false,' is false." But clearly the very same question can be raised about the referent of "This very statement" as it appears in the included sentence. And a similar question could be asked ad infinitum about any possible referent of "This very statement," since its referent would always include a referring expression, and the reference would always be to the referring expression just inasmuch as it is referring.¹¹

In short, the semantic paradoxes are due neither to the fact that they arise in self-referential statements nor to the fact that the self-reference is semantic. The precise difficulty is that there are no propositions expressed in these supposed statements; there is nothing definite to which the referring terms might refer. Like a mirage, the supposed referent continually recedes. Thus, until one sees that the semantic paradoxes are only putative statements, the semantic difficulties posed by such paradoxes cannot be removed.

Other self-referential statements, including the statements of semantically self-referential propositions, need not share these difficulties. "All propositions are either true or false" is semantically self-referring. "The proposition, 'All propositions are either true or false,' is either true or false" is an instance of

this proposition. This self-referential instance is not paradoxical; there is no difficulty in indicating its referent. Indeed, the proposition is true.

In the case of sententially and performatively self-referential statements, one does not find paradoxes formally like the logical and semantic paradoxes. Both the latter kinds of paradox are similar in form to the Liar Paradox, the statement of which—if it stated a proposition—would state a proposition which would be true if assumed to be false and false if assumed to be true. Performatively self-referential statements need not be paradoxical. For example, “I always write correct English” is not paradoxical.¹² Those performatively self-referential statements which might be called “paradoxical” can go wrong in a quite different way than logical or semantic paradoxes. “I never write correct English” might be called “paradoxical,” but it is not true if false and false if true. As soon as one considers whether the self-referential instance of this statement is true, one discovers that it is not. The sentence used to make the statement is evidence that the self-referential instance of the proposition is not true, and noticing this falsity ends one’s perplexity. The air of paradox here is due only to the fact that the proposition is falsified by the very sentence used to express it.

Also distinct from the logical and semantic paradoxes are performatively self-referential statements which go wrong performatively. The skeptic, for example, if he is to be consistent, must consider his own affirmation to be as groundless as all others. One cannot achieve anything by affirming a proposition which one admits to be groundless.

Nevertheless, performatively self-referential statements also can go wrong semantically. Such statements have difficulties similar to the difficulties of those semantically self-referential statements which cannot indicate a referent. For example, “This very affirmation is well-grounded,” which is performatively self-referential, has the same kind of difficulty as “This very statement is vague.” Since an affirmation is the affirming of a proposition, if one is to indicate the referent of the referring expression, “This very affirmation,” in the statement, “This very affirmation is well-founded,” then one must also be able to indicate the proposition which is affirmed. This, clearly, cannot be done.

However, in many cases it is neither impossible nor even difficult to indicate what is referred to in performatively self-referential statements. In performatively self-referential statements as well as in sententially self-referential statements the subject matter—that is, the act of affirming the proposition or uttering the sentence used in making the self-referential statement—can be identified by naming it or picking it out by a definite description.

In summary. There is nothing inherently illegitimate about statements which refer to themselves. It is now necessary to consider more systematically how those that go wrong do so.

C. How self-referential statements go wrong

Self-referential statements can go wrong in the same ways as other statements. The propositions expressed in such statements can be false or inconsistent with other propositions; the performances of such statements can be pointless. Sentences which seem to express self-referential statements can turn out to be meaningless.

The breakdown of self-referential statements insofar as they are self-referential is peculiar because some aspect of the statement itself, as referent of the proposition affirmed in the statement, gives rise to the falsity, meaninglessness, or pointlessness.

We have shown how some semantically self-referential statements can be such that it is impossible to indicate the propositions to which they allegedly refer. Since the very propositions which are supposed to be expressed in such statements are what they purport to refer to, these semantic difficulties justify the claim that such a putative statement expresses no proposition.

Semantically self-referential statements can also lead to performative difficulties. For example, "All propositions are false" is a pointless, semantically self-referential statement. It must be noted that "This proposition is false" is not an instance of the preceding general proposition. Its self-referential instance is "'All propositions are false' is false." Thus, if someone affirms "All propositions are false," his purpose is necessarily thwarted. He affirms a proposition which he must regard as false if he is to be consistent.

Finally, semantically self-referential statements can be falsified by their own propositions. We have shown that any statement of a general prohibition of self-reference is semantically self-referential. An instance of the proposition expressed in this prohibition refers to the proposition itself and says of it that it is non-self-referential. But this instance does refer to the proposition. Thus, the proposition is self-referential, and this fact falsifies the proposition. Likewise, assuming that propositions themselves are metaphysical entities, "No proposition can refer to a metaphysical entity" is falsified by its own proposition. Since this proposition refers to all propositions, it has an instance which refers to itself. Thus, the self-referential instance of the general proposition is falsified by the proposition itself.

Performatively and sententially self-referential statements also can go wrong in all of these ways. We have already shown that performatively self-referential statements can fail semantically by lacking reference. For example, "This very affirmation is well-founded" seems to be a performatively self-referential statement, but it can have no definite reference.

Performatively self-referential statements also can be performatively pointless; pointless statements of this sort are aptly called "self-defeating." We have already seen that skepticism is self-defeating in this way.¹³

The self-referential arguments for *Sfc* which we considered in chapter two, section E, are for the most part attempts to show that any statement of *Nfc* is self-defeating. If these arguments were successful, the affirmation of *Nfc* would be seen to be pointless when its self-referential instance is considered. James Jordan is very clear about this; he points out several times that his argument does *not* show that *Nfc* is false. He thinks it shows only that if *Nfc* is true, then there is no good reason to believe any thesis, including *Nfc*. Thus Jordan regards his argument as a practical, not a theoretical, argument for *Sfc*.¹⁴ The truth of *Nfc* remains possible.

Jordan is correct in thus limiting the claim he makes for his argument. In our own use of an argument of this form in chapter six, section G, we will take care to respect this limitation. If this kind of argument is successful, performative pointlessness is a characteristic of the affirming of *Nfc*, not of *Nfc* itself. Even if it is self-defeating for a *PNfc* to affirm *Nfc*, this does not show *Nfc* false. Likewise in the case of the skeptic: Even if his affirmation is groundless, it might still be true that all affirmations are groundless.

Because the effect of arguments similar to these two is to show the pointlessness of the opponent's act of affirming his thesis rather than to show the thesis false, arguments of this sort have been called "ad hominem." David Wiggins has thus characterized a standard self-referential argument against marxism.

The reply [that marxist beliefs also are conditioned] is no better than *ad hominem* because it leaves perfectly open the possibility that beliefs, capitalist, marxist, and all others, are *uniformly* tainted by the causality which determines them. It cannot tell against this that if it were so then nobody would have the knowledge of this fact but at best an accidental true belief. Perhaps that is how things are.¹⁵

Wiggins's criticism is correct to the extent that such arguments do not show any proposition to be false, but only show its affirmation to be self-defeating. Wiggins's criticism is not correct, however, to the extent that by calling such arguments "ad hominem" he suggests that they are necessarily fallacious. Such arguments are not ad hominem in the usual sense; they do not attack in some irrelevant respect the one affirming a position, but they attack the one affirming a position precisely insofar as he is affirming it. If one who uses this type of argument thinks he falsifies the position against which he is arguing, he is mistaken. For example, if those whom Wiggins criticizes suppose that their argument against marxism proves that theory false, they are mistaken. However, it is quite possible to employ this type of argument and fully recognize its limitations. Jordan does not think that his argument that the affirming of *Nfc* is self-defeating demonstrates the falsity of *Nfc*.

Performatively self-referential statements can also be *falsified* by the aspect of their performance to which they refer. We shall try to show in chapter six that if *Nfc* can be rationally affirmed—that is, if the attempt to rationally affirm it is

not self-defeating—then it is falsified by any rational affirmation of it, no matter how it is stated. For this reason, we next consider the falsification of performatively self-referential statements.

D. The falsification of self-referential statements

We define the falsification of a performatively self-referential statement as follows: A self-referential statement (S) in which there is an affirming (A) of a proposition (P)—where P is not a logical truth—by an uttering (U) of a token (T) will be performatively falsified if and only if there is a property (Q) such that either A or U has Q and “ A has Q ” or “ U has Q ” is inconsistent with P . For example, the stating of “This statement is not printed on paper” is falsified by its being printed here. The method of expressing the statement has a property—being printed on paper—such that the statement that this expression has this property is inconsistent with the proposition stated. Of course, this proposition would not be falsified by the way of expressing it if it were spoken.

In addition to performatively self-referential statements there are performatively self-referential *propositions*; such propositions are performatively self-referential no matter how they are performed. Some of these propositions are such that they are falsified by their performance and would be falsified by *any* performance of them. A performatively self-referential proposition (P) is performatively falsified if and only if there is a property (Q) such that *any* affirmation (A) of P or any utterance (U) of any token (T) used to express P has Q and “ A has Q ” or “ U has Q ” is inconsistent with P . For example, an instance of the general proposition expressed by the statement (S), “No utterance can be used to express a proposition,” is performatively self-referential. It refers to the utterance used to make S . This instance is falsified by this utterance. This utterance has a property—of expressing a proposition—such that its having this property is inconsistent with the performatively self-referential instance of the general proposition affirmed in S . Moreover, it is clear that this proposition will have a self-referential instance in any statement of it and that this instance will be falsified by the utterance in any statement in which the proposition is stated.

The falsity of performatively self-referential propositions which are performatively falsified can be philosophically important. If a philosophical theory goes wrong in this way, the theory should be rejected *as false*. The proposition affirmed in any statement of the theory is falsified by any performance in which it is stated. Thus, if a philosophical theory has a performatively self-referential instance, and if that instance is falsified by the performance of the statement, or, to be more precise, is such that it would be falsified by any performance of stating the theory, then the theory is inevitably falsified.

We call philosophical arguments which show a theory to have self-referential

difficulties “self-referential arguments.” In some cases, the self-referential instance is shown by the argument to be inevitably falsified. In others, the self-referential instance is shown by the argument to render self-defeating any statement of the theory.

Descartes’s *cogito* can be construed as a self-referential argument.¹⁶ “I do not exist” is falsified by its own affirmation, since affirmations and all other propositional acts have the property of being made by someone who does exist.

The well-known argument that the statement of the verifiability criterion is self-refuting can also be understood as a self-referential argument. The verifiability criterion can be construed as the claim that there are more utterances than there are utterances expressing propositions. It is a proposal of a criterion to distinguish these. Thus, its self-reference will be performative; it refers to utterances as purported statements.

Assuming, for the sake of this example, that the statement of the verifiability criterion affirms a proposition, we state the self-referential argument against the verifiability criterion as follows:

1) The verifiability criterion is the proposition (*V*) which states: Any sentence (*S*) which expresses a proposition necessarily has the property (*Q*)—that is, the property of expressing an analytical truth or an empirical hypothesis.

2) Any statement of *V* is a case of *S*, and a statement of *V* will express a proposition if and only if it has *Q*.

3) Any statement of *V* lacks *Q*. (The statement of *V* must lack *Q* in order to perform its function of excluding as meaningless all sentences which lack *Q*. In other words, the statement of *V* must lack *Q* if the statement is to achieve its purpose. If it has *Q* then it will be pointless. If the statement of *V* is an empirical generalization about utterances which have been discovered to be meaningless, it cannot exclude the possibility that meaningful utterances not having *Q* will be discovered. If the statement of *V* is a stipulation or definition, it cannot exclude the possibility that there are sentences which are meaningful by some other definition of meaning.)

4) It is not the case that any sentence used to express *V* expresses a proposition.

5) But *V* is a proposition.

6) Any sentence used to express *V* expresses a proposition.

(1) is a statement of the verifiability criterion. (2) is a statement of its self-referential instance. (3) is a description of the statement of *V* which is required for this statement to be a purposeful statement. (4) is entailed by (2) and (3); it is what the verificationist must say about his own sentence if he is to remain consistent and to achieve the purpose of his statement. (5) is true by assumption. (6) states what is true of the sentence used to express *V* on this assumption. (6) and (4) are contradictories. (6) states a fact about the sentence used to express *V*; (4) states what follows from *V* together with the conditions for the

purposeful affirming of *V*. (6) falsifies (4). Any statement of *V* is thus falsified by its own performance.

In establishing inevitable falsity, self-referential arguments of the kind exhibited in the two previous examples are different from self-referential arguments by which someone seeks to show that a philosophical statement is self-defeating. The latter kind of argument does not demonstrate the falsity of the proposition in question, even though it can show that any act of affirming the proposition is inevitably self-defeating. Arguments by which someone seeks to show the falsity of a performatively self-referential proposition bear, not merely upon the affirming of it, but upon its truth.

There are several objections to self-referential argumentation. First of all, someone might object that self-referential arguments leave open the possibility that the opposed thesis could be true. This objection might be based upon a confusion between self-referential arguments which reveal falsity and self-referential arguments which show a statement to be self-defeating.

This objection might also be based on a confusion between logical impossibility and the inevitable falsity which certain self-referential arguments reveal. The proposition refuted by a self-referential argument which reveals the self-referential falsity of it is logically coherent; it picks out a possible state of affairs. The objection we are considering might be based on the assumption that since the proposition picks out a possible state of affairs, it remains possible that the proposition be true. One knows, however, that this state of affairs does not obtain, because the proposition is falsified by its own performance. In fact, where a performatively self-referential *proposition*—in contrast with statement—is falsified, one knows that the falsity is inevitable. There is no way in which the proposition can be stated which does not also provide the falsification of the proposition. Still it is not logically impossible that the state of affairs obtain.

In other words, performatively self-referential propositions which are falsified are not logically impossible nor are their contradictories logically necessary truths. Their falsity is in some respects like that of a falsified scientific theory; one does not regard such falsity as indicating a mere contingency, but rather as revealing a kind of necessity about the world. One expects that the theory will be falsified in all instances which are essentially like those that first falsified it. A scientific theory does not merely *happen* to be false.

Likewise, performatively self-referential propositions which are falsified by their own performance do not merely *happen* to be false. The falsity reveals a kind of necessity. However, unlike scientific theories, performatively self-referential propositions which are false carry with them, in their very statement, the fact which falsifies them. The falsification is inevitable.

Another possible objection to self-referential arguments is that there is a simple and effective way to blunt their force. A philosopher whose thesis is

criticized by a self-referential argument can avoid the force of the criticism by limiting the scope of his thesis so that the self-referential instance does not arise.¹⁷ Thus, the self-referential argument does not terminate the philosophical controversy.

But to limit the scope of a thesis which has been refuted by a self-referential argument is implicitly to admit that the original thesis was indefensible. The claim is no longer the same. Perhaps other instances of the thesis are true. But by this limitation the self-referential instance is admitted to be indefensible. The fact that the philosophical discussion continues in no way shows that the self-referential argument has not been decisive. The discussion can go on but with a difference; the thesis originally affirmed has been admitted to be indefensible. Both parties to the discussion can continue the controversy. The critic can ask by what principle the self-referential instance is abandoned although others are still claimed to be true. To abandon the self-referential instance because it has been shown indefensible while still claiming that the thesis holds in other instances is to make an arbitrary move, unless one provides a basis for distinguishing the falsified instance from the others. If one provides such a basis, his new position is clearly different from the original thesis he affirmed.

Of course, neither of the preceding objections against self-referential arguments disputes the claim that the performatively self-referential proposition at issue is indefensible. But one can dispute this claim. Self-referential arguments can beg the question. They can do this by ascribing to the performance a property which one who affirms the proposition need not ascribe to it.

The examination and criticism in chapter two, section E, of previous attempts at a self-referential argument against *Nfc* revealed that these attempts are question-begging. The *PNfc* can legitimately answer these arguments by pointing out that they assume what he need not admit—namely, the incompatibility between something's being a reason for belief and its being wholly determined by causal conditions.

However, the fact that some self-referential arguments are question-begging does not mean that all are. In fact, there are some cases in which ascribing a property to the performance which renders the statement indefensible cannot be question-begging. This will be the case whenever the one making the performatively self-referential statement must either regard his performance as having the property in question or become inconsistent.

A proponent of the verifiability criterion might object that it is question-begging to assume—as we do in our criticism of it—that the verifiability criterion is a proposition. He might say that the verifiability criterion is a rule of meaning, not a proposition. In saying this, he could be correct; it is possible that his utterance does not express a proposition; it might be significant in some

noncognitive way. It might, for example, express the feelings of verificationists about metaphysics or it might be an exhortation. However, the verificationist could wish to make some claim for the truth or adequacy of his position. As we have shown, if he does affirm his position, then no matter how he affirms it, he makes some sort of truth-claim for it, and in making this claim he states a proposition. Of course, his statement will not meet his own criterion for stating a proposition, but he nevertheless will make a claim for something. In this sense he states a proposition. If he does not admit this, his utterance has no cognitive force, and thus he cannot thereby deny the cognitive meaningfulness of utterances which do not meet his criterion. In short, the verificationist must admit that his thesis is a proposition if he is to achieve his purpose. Nor can the verificationist, without defeating his purpose, admit that his statement is either an empirical hypothesis or a definition.

Our argument against verificationism stands in contrast to the attempted self-referential arguments which we criticized in chapter two, section E, against the affirmation of *Nfc*. In that case, the *PNfc* was able to deny without rendering his argument pointless that the performance of his statement has the falsifying property. He was able to explain that although his affirming of *Nfc* is itself causally determined it can nevertheless be rationally performed. Thus, the only effect of those self-referential arguments was to elicit a clarification of the theory criticized.

Since many attempts to demonstrate *Sfc* have been question-begging, we shall be especially careful to avoid this fallacy. In formulating our argument against *Nfc*, we shall try to show—in terms of forms of argumentation clarified in the present chapter—that if *Nfc* can be rationally affirmed, then it is performatively falsified, and that if it cannot be rationally affirmed, then any attempt to affirm it will be inevitably self-defeating.

Another possible objection to the force of self-referential arguments is that even if the propositions which they show to be false are inevitably falsified by their own performance, they *might*, nevertheless, be important truths which, unfortunately, cannot be stated.

There is something very strange about the notion of “important truths which, *unfortunately*, cannot be stated.” The difficulty which arises in stating propositions which are inevitably self-defeating or falsified is not a mere misfortune—that is, an avoidable accident. The falsity or self-defeating character of such propositions is established by *any* stating of them. They can be stated; they are stated; they are falsified or they defeat their own purpose in being stated. To say that they cannot be stated is merely to say that whenever they are stated, they are falsified or shown to be pointless.

We do not deny that there might be a true proposition which cannot be stated. But if there is any such proposition, clearly it cannot be one such that if it could

be stated, it would be inevitably falsified by its own affirmation. Any such proposition is false. It *might* be true only in the sense that it is not logically false.

When the objector says in respect to propositions which are self-referentially falsified, "They *might*, nevertheless, be important truths," this claim can be taken in either of two senses.

In one sense, it means that the propositions shown to be inevitably false by their own performance are logically possible. To say that a proposition is logically possible is to say that it could be either true or false. To be able to be either true or false is just to have propositional sense and reference. Therefore, a proposition which could not be true in this sense is logically impossible. It follows that if the objector means his objection in this sense, then when he says "They *might*, nevertheless, be important truths" what he says is equivalent to "They might, nevertheless, be important falsehoods."

In another sense, "They might, nevertheless, be important truths" could mean something other than that self-referentially falsified propositions are logically possible, that is, neither necessary nor impossible. In this other sense, a preference for the truth of falsified propositions is expressed. Such a preference can hardly be rational.

Someone might accept the preceding response with respect to propositions which are self-referentially falsified, but object that if the proposition is only shown by a self-referential argument to be performatively self-defeating, it might—in a more significant sense—still be true and important. We admit that such propositions might be true in a more significant sense than that in which a falsified proposition might be true. A falsified proposition is merely logically possible; a proposition shown to be self-defeating is not falsified, and there remains an ontological possibility that it be true.

Nevertheless, a proposition shown by a self-referential argument to be performatively self-defeating cannot be used as a premise or defended as a conclusion in a rational discussion. Rational discussion involves both propositional acts such as affirming and speech acts such as uttering. In our argument in chapter six, nothing turns upon reference to utterances. Thus we reply here to the objection only insofar as it applies to self-defeat arising from reference to the act of affirming.

To claim a proposition true is to affirm it; to rely in reasoning upon the fact that a proposition has not been falsified is to weakly affirm it. There can be no point in affirming, even weakly, a proposition whose affirmation is inevitably self-defeating. It must be noted, moreover, that rational discussion can occur in an individual's own thinking. Thus, if it is pointless for a person to affirm a proposition in an argument with another, it is pointless for him to think the proposition true in his heart, while avoiding embarrassment by not uttering it aloud.

E. The affirmation of “No free choice”

If *Nfc* is true, then no one’s affirmations and utterances can be freely chosen acts. Affirming and uttering certainly are among a person’s acts, but if *Nfc* is true, no one can freely choose to do these acts any more than he can freely choose to do any other acts. If no one’s affirmations and utterances can be freely chosen acts, then no *PNfc*’s affirming or uttering can be a freely chosen act. The *PNfc* is a man among men; if no one can make a free choice, neither can he. If no *PNfc*’s affirming or uttering can be a freely chosen act, then no *PNfc*’s stating of *Nfc* can involve a freely chosen act. Whenever a *PNfc* states his thesis, he performs an instance of the kinds of acts which are affirmations and utterances. Thus, if *Nfc* is true, any statement of it, in excluding the possibility of free choice generally, excludes free choice from the very act which affirms it.

A proposition’s truth or falsity depends upon whether the state of affairs it picks out obtains. Whether or not the state of affairs a proposition picks out obtains, the reference of the proposition remains the same. Thus, the reference of *Nfc* remains the same whether it is true or false. *Nfc* has an instance referring to any possible human act; thus, it has an instance referring to any possible act of affirming or uttering, and so to any possible act of affirming *Nfc* and uttering sentences which express *Nfc*. Hence, it is clear that *Nfc* is a performatively self-referential proposition.

We argue in chapter six that *Nfc* is self-refuting—that any affirmation of *Nfc* relevant to *Sfc/Nfc* either falsifies *Nfc* or renders the affirming of it self-defeating. To prepare for this argument, we now clarify what is involved in the affirming of *Nfc*. To do so, we must clarify what is involved in making any grounded affirmation.

At the beginning of this chapter, we defined an “act of affirming” as any propositional act by which someone holds a proposition to be true or reasonable to accept. Assenting to a proposition by faith, accepting it as a hypothesis which has some likelihood, and asserting it are among the ways of affirming it. Thus, one affirms a proposition whenever he holds it to be true or more likely to be so than not.

“To affirm” sometimes is taken to mean to publicly assert a proposition as absolutely certain. In this sense of “to affirm,” its contrary is “to disavow”; one only denies a proposition in this sense if one publicly rejects the proposition as certainly false. Between affirming and denying in these strong senses, there obviously are many propositional acts which involve holding a proposition true or likely. By contrast, according to the definition of “affirming” which we have adopted, any propositional act by which one holds a proposition to be true or reasonable to accept is an affirmation of the proposition.

One can consider any proposition either by itself or in relation to other propositions to which it is related truth-functionally. One also can perform propositional acts which have nothing to do with the truth or falsity of a proposition—for example, one can include a proposition in a fictional narrative. We are not interested at the moment in propositional acts which bear upon propositions in truth-functional relations, and we are interested only in propositional acts which bear upon a proposition's truth or falsity. Considering a proposition by itself and with respect to its truth or falsity, one is capable of only three kinds of propositional acts. First, one can merely entertain the proposition as one whose truth or falsity—or even likely truth or falsity—is yet to be discerned; second, one can hold the proposition true or regard it as more likely true than not; third, one can hold the proposition false or regard it as more likely false than not. The second kind of propositional act is what we call “affirming”; the third, what we call “denying.” Obviously, on these definitions, one never affirms or denies without doing both, for whatever proposition one affirms, one denies its contradictory.

Thus, even if one only thinks privately that a proposition might be true, he affirms it. If one says to himself, “I am inclined to think that p ,” he affirms p ; perhaps p is a most tentative belief which he would not wish to mention to anyone else, yet he affirms p . Expressions such as “Don’t you think that p might be true?” and “I take it that p cannot easily be denied” and “It is my personal feeling that p ” and “Isn’t it perhaps more realistic to think p ” and “No one today can take $not-p$ very seriously” and “I think it clear that the time has come to rethink the traditional view that $not-p$ ” can be and often are used to affirm p .

It follows that one affirms if and only if he also denies. Thus, if one expresses a possible belief but admits that the contradictory is equally likely to be true, he is not affirming the belief. Sometimes a hypothesis is put forward without being affirmed; hypotheses can be proposed as questions or included in the design of a research project without any judgment being made with respect to their truth. On some theories of practical and theoretical procedure, it is even possible to make use of hypotheses while admitting that their contradictories are equally likely to be true. In such uses of hypotheses to control the environment or to organize a region of experience, the hypothesis is not proposed for acceptance as true or more likely, and its contradictory is not regarded as any less reasonable to suppose true.¹⁸ Thus, expressions of possible beliefs and proposals of hypotheses can occur without any propositional act of affirming. In such cases, the proposition is proposed with no more claim about what is the case than would be made if one articulated the propositions in a purely fictional account.

For the sake of clarity, we now make several points which we do not expect any *PNfc* to dispute.

In *Sfc/Nfc*, it is clear that the *PNfc* must affirm *Nfc*. If one merely proposes the proposition as an interesting possibility, neither affirming it nor denying it, he is not a *PNfc*, for his propositional act in respect to the proposition is no more that of a *PNfc* than that of a *PSfc*. In a dispute about which of two contradictory propositions is true, one cannot be located on either side unless one expresses at least very tentatively a propositional act leaning toward one side.

To put the same point in other words. *Nfc* describes a world from which the ability to make free choices is absent. If any such world exists, then in that world no one can make a free choice. But the mere proposition, *Nfc*, does not claim that the actual world excludes *Sfc*. This claim is made only when someone holds *Nfc* true or likely true of the actual world—in other words, only when someone affirms *Nfc*. The mere proposition, *Nfc*, can be doubted or denied as well as affirmed. One who doubts it or denies it does not exclude free choice from the actual world, but leaves room for it.

Affirmations can be made in different ways; they can be groundless or they can be grounded.

A groundless affirmation is a gratuitous preference for one of a pair of contradictories. One of the pair is held without any epistemic warrant to be true or more likely. For example, a person knowing nothing about space exploration might affirm that travel to distant stars will be achieved within the next century, and another equally ignorant person might dispute this by saying that such travel is impossible. Of such groundless affirmations and denials, barroom arguments are made. One also can make an affirmation despite the availability to him of grounds for making the contradictory affirmation. For example, a person who is told that he is dying of incurable cancer might affirm for many months that he is not seriously ill and that he surely does not have cancer.

In such cases, the groundless affirmation and the irrational affirmation remain affirmations. One proposition is held to be true or likely; its contradictory is held to be false or unlikely. A person who publicly affirms a proposition without grounds asks others to accept it and to reject its contradictory, but if anyone asks why the proposition should be accepted, no reason can be given for accepting it. The reasonableness of accepting it, or even of considering it, is not shown.

If the *PNfc* affirmed his position in this way, he would exclude *Sfc* only by an ipse dixit. However, *Sfc/Nfc* is a controversy to which the parties are scientists, philosophers, theologians, and others who are, or who at least claim to be, engaged in the serious pursuit of truth. In such company, anyone who makes arbitrary affirmations is ignored as soon as the arbitrariness of his affirmations becomes clear. Thus, the *PNfc* not only affirms *Nfc* but also proposes it as a grounded affirmation.

Affirmations can be grounded in different ways: by direct evidence, by

logical insight or analysis, and by other procedures—for example, inductive argument.

If one has direct evidence for a proposition, one's affirming of it will be epistemically legitimate. For example, if one sees rain falling, one's affirmation that rain is falling has epistemic legitimacy. In such a case, the grounds for the affirmation are immediately given. If one has evidence of this sort, it is reasonable to affirm the proposition for which one has it, but it is not merely more reasonable to affirm the proposition than its contradictory. It is unreasonable and perhaps impossible to deny a proposition for which one has such evidence.

If one understands the formula of a logically true proposition, one's affirmation of it also will be epistemically legitimate. For example, if one understands $6 + 7 = 17 - 4$, then one's affirming of the equation has epistemic legitimacy. The same can be said for one's affirming of the conclusion of a more complex logical or mathematical proof. Here again the claim is not that it is more likely that the proposition is true or more reasonable to accept, but rather that it is perhaps impossible not to assent to it.

Statements of fact based on immediate evidence and statements of logically true propositions are not the only sorts of grounded affirmations. There are many statements—for example, interpretations of data, generalizations, and hypotheses—which one reasonably affirms, although they are based only somewhat indirectly on evidence. There are also propositions which are shown to be true by the analysis of language or by conceptual clarification which are reasonably affirmed, although they are neither self-evident nor logically necessary. There also are propositions supported by authority which one reasonably affirms. A great many—if not most—of the truths people think they know are propositions which they affirm on the authority of parents, teachers, friends, neighbors, journalists, technical experts, scientists, religious leaders, and so forth.

We define as "rational affirmations" all affirmations which are grounded otherwise than by direct evidence or by insight into logical truth. One affirms rationally if and only if the proposition he affirms is one more reasonable for him to hold true or likely than its contradictory. The contradictory of a rational affirmation remains consistent both with the direct evidence one has and with the logically necessary propositions one knows.

In the present work, we are concerned with a controversy regarding a set of facts—the phenomena of choice. In chapters two and three we pointed out that some on each side of the controversy have suggested that the affirmation of their position is based upon direct evidence. But neither of these claims succeeds; both depend upon a failure to recognize the limits of what can be derived from the data alone. Nor does an appeal to logically necessary truths settle the controversy. A clear example of this is the failure of the fatalist

attempt to derive *Nfc* from logical truths alone. It remains that the affirmation either of *Sfc* or of *Nfc* can be at best a rational affirmation.

Thus, as we have seen, *Nfc* is affirmed as a hypothesis by physical and psychological determinists. Some psychological determinists and some who attack the intelligibility of free choice seem to regard clarification of the phenomena of choice or conceptual analysis as an important ground for their affirmation of *Nfc*. Operationalists consider their affirmation of *Nfc* to be justified rationally by its fruitfulness. Religious believers who affirm *Nfc* think their assent to it is justified by the authority of their faith.

In chapter six, we argue that any attempted rational affirmation of *Nfc* either falsifies *Nfc* or renders the attempt to affirm it self-defeating. Our argument will rest on the claim that there are necessary conditions for rationally affirming *Nfc*—that is, conditions which must be fulfilled if the attempt to rationally affirm it is to succeed. To clarify the notion of *conditions for rational affirmation*, we note that many human acts have conditions which must be met if these acts are to be what they are intended to be.

As we mentioned previously, C. K. Grant points out that making a statement has what he calls “pragmatic implications,” and that if these implications are false, the making of the statement is irrational. Somewhat similarly, John R. Searle explains that an illocutionary act such as promising has a set of conditions which must be satisfied for a successful and nondefective act.¹⁹

In law, conditions are laid down for the success of acts such as making a contract or a will, indicting a person, passing sentence on a convict, and so on. If the conditions are not met, the attempted legal act can be held null and void. In religion, likewise, there are conditions for the validity as well as for the licitness of ritual acts, such as sacraments; the ritual is believed to be pointless if conditions for its validity, at least, are not satisfied.

It seems reasonable to suppose that there are analogous conditions for the propositional acts of affirming and denying. If there are such conditions and they are not fulfilled, then acts of affirming and denying will be defective in some way. Such conditions for rationally affirming a proposition would be neither a function of the meaning or truth of the proposition nor of the meaningfulness of the language used in stating it.

Assuming there are conditions for making rationally grounded affirmations, then if these conditions are not fulfilled an attempt to make such an affirmation fails to be what it was meant to be—a rational affirmation. Of course, one’s affirming can succeed in this respect—when these conditions obtain—and one’s statement can still go wrong in some other way. For example, one can succeed in affirming a false proposition or a proposition inconsistent with other propositions of which one is more certain. One might have rational grounds for affirming a proposition—for example, one might have evidence which confirms a plausible hypothesis—yet the evidence itself might consist in faulty

observations. For these reasons, an attempt to make a rational affirmation can succeed, although the proposition one affirms is false, inconsistent with other propositions which one should prefer to it, or actually without the ground one supposes one has.

Thus, it seems that certain conditions must obtain if one is to be successful in rationally affirming a proposition such as *Nfc*. Inasmuch as we regard *Sfc* as true, we obviously consider *Nfc* false and think there are many propositions inconsistent with *Nfc* which one should prefer to it. Moreover, in chapter three, we have disputed the soundness of all the grounds we know of on which *Nfc* has been affirmed. However, we have not disputed—in fact, we have made clear—that *Nfc* is rationally affirmed. Thus, we assume that the conditions for rationally affirming *Nfc* can obtain. We next consider certain of these conditions.

F. Rationality norms as conditions of affirmation

In each of the sections of chapter three, we examined one of the main lines of argumentation for *Nfc*. We ended each section with a summary of the *PNfc*'s argument. These summaries included explicit articulations of certain principles which the various arguments presuppose.

The following principles are excerpted from the summaries. A full description of the data is to be preferred to a partial description. A generalization based on meticulous observation is to be accepted. Logical principles are to be adhered to in all one's thinking, even when doing so requires one to give up beliefs based upon experience. Any view which meets the criteria of simplicity, predictive success, and explanatory power . . . is to be accepted. Relevant facts are not to be ignored in a theoretical dispute. A method of interpretation which is successful is to be relied upon in further, similar cases. An account of phenomena is not to be accepted if it requires something inconsistent with the data for which it is supposed to account. Nothing inexplicable is to be admitted as possible. If one accepts a goal, a view of things helpful for achieving that goal is to be preferred—other things being equal—to a view of things which blocks effective pursuit of it. If one is reasonable in accepting traditional faith, then all its doctrines, without qualification, are to be accepted as meaningful and true.

We maintain that these principles—or some such principles—are necessary conditions for rationally affirming *Nfc*. We call principles of this sort “rationality norms.” While the rationality norms used in arguments for *Nfc* might be refined, we do not maintain that the *PNfc* is unreasonable in appealing to these or similar principles. We do maintain that the rationality norm or norms which a *PNfc* presupposes in any argument for his position must be in force, not null, if

the act of rationally affirming *Nfc* is to succeed—that is, if it is to be a valid and not a mere putative rational affirmation.

Rationality norms guide one in affirming propositions. They bear directly on the affirming, not on the proposition affirmed. They do not describe facts which might confirm the proposition, nor are they generalizations from which the proposition might be deduced. Rationality norms are *norms*; they say what *is to be*, not what *is*. They direct one's steps in moving toward making an affirmation and in making it, so that one's affirming will be grounded, even though the proposition one affirms is neither logically true nor a matter of evident fact. In rationally affirming a proposition, one assumes rationality norms, not as premises from which one might deduce conclusions about the world, but as licenses or warrants legitimating the moves one makes in taking one proposition rather than its contradictory as more likely to be true of the world.

In one respect, rationality norms are like the rules of formal logic. Both bear upon the legitimacy of the relationship between grounds and a conclusion; rationality norms are concerned with the relationship of evidence or reasons to a proposition which they rationally support somewhat as rules of formal logic are concerned with the relationship of premises to a conclusion which they entail. Neither a rationality norm nor a logical rule such as *modus ponens* is any part of what is affirmed in the proposition on whose affirmation it bears.

However, the rules of formal logic differ from rationality norms. The rules of formal logic show how propositions and formally distinct parts of propositions are formally related, while rationality norms indicate only that certain sorts of reasons and evidence provide adequate grounds for affirming a conclusion.

Of course, *Nfc* can be deduced from some more general propositions. For example, universal determinism entails *Nfc*. But universal determinism itself gains its plausibility from a scientific theory of physical determinism; as we shall explain more fully shortly, the affirmation of such a theory presupposes a rationality norm—for instance, a simplicity-rule. Even if *Nfc* could be deduced from the laws of logic, as the fatalist attempts to do, the affirmation of *Nfc* on such a basis would require justification, because such an affirmation would involve a preference for logical truths over other grounds for belief.

Another consideration further clarifies the distinction between the rules of formal logic and rationality norms, and the bearing of the latter upon the *act* of affirming. One who violates a law of logic eventually finds himself in formal incoherence; he loses his ability to *mean* anything and to *say* anything. One who violates a rationality norm does not thereby become formally incoherent. If the violation is clear, he will be unreasonable, but he can still talk in a way which makes *sense*. His propositions are coherent, but his affirming of them is

somehow out of order. He can be called “foolish,” “rash,” “careless,” “dim-witted,” “silly,” or something of the sort.

Thus rationality norms bear upon the act of affirming; one who follows them makes affirmations rationally, while one who ignores or violates rationality norms proceeds unreasonably in making affirmations.

Thus, we claim that there are nonformal norms of affirmation, presupposed by the *PNfc*'s affirming of *Nfc*. This claim is not ad hoc, and the *PNfc* is by no means alone in presupposing such nonformal norms.

Similar norms also function in the arguments for *Sfc* which we criticized in chapter two. The argument for *Sfc* based upon experience proceeds on the principle that if something seems to be so, then it is to be taken to be as it seems. The argument from moral responsibility proceeds on the principle that the presuppositions of commonly held beliefs and important institutions are to be assumed to obtain. The argument proposed by Thomas Aquinas assumes that a position which is consonant with one's entire worldview is to be accepted. William James's argument supposes that if belief in a position contributes to the moral quality of life, then that position is to be preferred to one which detracts from such quality. The argument that *Nfc* is self-defeating assumes that any self-defeating position is to be given up.

Moreover, our own argument in chapter six, section I, insofar as we attempt to establish *Sfc*, appeals to certain rationality norms.

Many philosophers in recent times have articulated principles for inquiry and rational affirmation, and pointed out that these principles—to which they have given various names—are neither descriptive statements nor formal logical rules.

Wilfrid Sellars, for example, emphasizes that anyone who engages in conceptual activity must recognize norms and standards:

. . . if one gives to “practical” the specific meaning *ethical* then a fairly sharp separation of these activities can be maintained. But if one means by “practical” *pertaining to norms*, then so-called theoretical reason is as larded with the practical as is practical reasoning itself.²⁰

Sellars makes clear that the epistemic *values* of theoretical reasoning bring it into close relationship with practical reason. A full theory of practical reason “would also recognize the inseparability, yet distinguishability, of theoretical and practical reason in all dimensions of human life.”²¹

Roderick Chisholm also emphasizes the normative dimension of epistemology. Throughout his writings in this field—a field which many regard as wholly theoretical—he presses the analogy between epistemic and ethical judgments. Some of the epistemic principles he formulates are very similar to some of the rationality norms we have listed. He also speaks of “epistemic obligations” and talks of a proposition's being “worthy of belief.” Chisholm

clearly thinks that the moves governed by such norms are neither standard deductions nor inductions, and that the norms themselves are neither descriptive statements nor mere conventions.²²

Other philosophers who do not speak as explicitly as Sellars and Chisholm do about the normative principles of theoretical inquiry and affirmation nevertheless sometimes mention such norms. Quine, for example, formulates an empiricist rule about the leaps which are to be taken from what is given: "Don't venture farther from the sensory evidence than you need to."²³ Arthur Murphy points out that a theorist should avoid infatuation with his theory if his theorizing is to be "a serious activity, responsibly carried on, and subject to practically normative standards for its right performance."²⁴ Max Black, in explaining why one should accept the conclusions of inductive arguments, points out that human beings belong to the "inductive institution" and are thus subject to "norms of belief and conduct imposed by the institution."²⁵

We hold that *Nfc* could not be rationally affirmed without implicitly assuming if not explicitly invoking some rationality norm or other. If we are right, it follows that any *PNfc*'s affirmation of his position is conditioned by at least one rationality norm. Several considerations indicate the correctness of this contention.

First, the ways in which *Nfc* is affirmed in *Sfc/Nfc* make clear the essential role of rationality norms in the affirming of it.

Often *Nfc* is affirmed as part of a larger theory or worldview. Such a theory or worldview gains whatever credibility it has insofar as it is a hypothesis which plausibly accounts for a certain range of data, or insofar as it is an interpretation of experience as a whole. However, there are many competing theories and worldviews, and a person prefers one to another only insofar as one is simpler, more far-reaching, more pleasing, more useful, or something of the sort. But preference by such criteria assumes that a hypothesis or interpretation of experience which meets them *is to be preferred* to one which fails to meet them or meets them less well.

As we explained in chapter three, sections A and C, it is not surprising that *Nfc* comes to be affirmed in the context of a general theory or worldview. It is a universal, negative proposition about a certain conceivable human capacity; it says that no one has that capacity. Such a proposition is not simply descriptive; it is not merely a generalization; and, since it is a claim about the world, it is not merely a logical truth. A claim such as this needs some sort of indirect justification, and one way to give it such justification is by fitting the proposition to be affirmed into a special or general theory which seems reasonable insofar as it meets the criteria established by some rationality norm.

This point can be made more specific by considering the affirmation of *Nfc* as part of the hypothesis of physical determinism, and by seeing the role of a rule of simplicity in the affirmation of such a hypothesis.

If a proposition is affirmed as a hypothesis, alternatives to it remain possible; its contradictory is logically possible and its contradictory can admit the very same data, for any hypothesis goes beyond immediate evidence. Thus, logic and immediate evidence alone do not compel one's assent to a hypothesis.

The purpose of a hypothesis is to account for the data. The only justification for admitting factors into a hypothesis is that they serve this purpose. Moreover, if a theorist admits any unnecessary factors into his hypothesis, he assumes an unnecessary liability. From these considerations it follows that—other things being equal—a more economical account of the data is to be preferred. Of course, there can be arguments about which of two or more hypotheses is more economical. Moreover, economy is not the only desideratum in inquiry; someone can argue for a less economical account if it better satisfies some other criterion. Still, at least some degree of conformity to a rule of simplicity is necessary for a proposition if it is to be rationally affirmed as a hypothesis.

Clearly, when a physical determinist affirms *Nfc* as part of his general hypothesis, he implicitly appeals to a simplicity rule.²⁶ The phenomena of choice do provide some ground for judging that one has made a free choice. However, the physical determinist supposes that admitting such an ability would complicate unnecessarily his coherent view of nature. The citadel of will would be holding out after everything else had surrendered to deterministic theory. In this case, a deterministic account is preferred because it excludes a property—the ability to make free choices—which would be diverse in kind from any other natural property. This peculiar property is not directly given in experience. It seems reasonable to hope that man's physical and biological functions can account for all human behavior, rendering unnecessary the supposition of so odd a property. The appeal to a rule of simplicity is seldom made explicit in so many words. But the appeal clearly is implicit in the arguments of the physical determinist.

Thus, it is clear that even if a *PNfc*'s version of physical determinism were to vary more or less from the typical examples of it we considered in chapter three, section C, any version of this position would have to appeal to some sort of simplicity rule. In some cases, the simplicity rule might be assumed into a demand for predictive success or fruitfulness in practical application.

We also considered in chapter three, sections D and E, arguments for *Nfc* based upon a clarification of the role of purposes in choice and upon the inexplicability or mysteriousness of choices if they are free. Such arguments need not appeal to a simplicity rule, for they do not account for the data in the same way a hypothesis does. But arguments for *Nfc* on these other bases also attempt to rule out *Sfc* as oversimplified, incomplete, or unintelligible. As we saw, all such criticisms involve the application of criteria of preference for *Nfc* and an implicit appeal to the norms which demand that these criteria be met.

Many of the arguments for *Nfc* depend upon an implicit appeal to the principle of sufficient reason. Historically, this principle often has been thought of as an a priori truth, a basic law of being and of knowledge. However, Kant was on the right track, we think, in regarding it as a regulative principle. It does not function in arguments otherwise than as a rationality norm. Richard Taylor, who accepts the principle, treats it as a “datum—not something which is provably true, but as something which all men, whether they ever reflect upon it or not, seem more or less to presuppose.”²⁷ It seems clear that the principle of sufficient reason makes a normative demand: An adequate reason why anything is so rather than otherwise *is to be expected*. If this demand is qualified by an additional phase—*unless one has a reason not to expect such a reason*—the principle of sufficient reason loses its metaphysical ring but becomes a plausible norm for rational inquiry and affirmation.

The preceding clarification of the conditions for affirming *Nfc*, whether as a hypothesis or not, indicates that some rationality norm or other must be presupposed in any rational affirmation of this position. We doubt that any *PNfc* would deny this, for such norms have a pervasive role in inquiry. They are required for what Wilfrid Sellars calls the “material moves” which are necessary in any scientific language:

Everyone would admit that the notion of a language which enables one to state matters of fact but does not permit argument, explanation, in short *reason-giving*, in accordance with the principles of *formal logic*, is a chimera. It is essential to the understanding of scientific reasoning to realize that the notion of a language which enables one to state empirical matters of fact but contains no material moves is equally chimerical. The classical “fiction” of an inductive leap which takes its point of departure from an observation base undefiled by any notion as to how things hang together is not a fiction but an absurdity. The problem is not “Is it reasonable to include material moves in our language?” but rather “Which material moves is it reasonable to include?”

Thus, there is no such thing as a problem of induction if one means by this a problem of how to justify the leap from the safe ground of the mere description of particular situations, to the problematical heights of asserting lawlike sentences and offering explanations. The sceptics’ notion that any move beyond a language which provides only for the tautologous transformation of observation statements is a “venture of faith” is sheer nonsense. An understanding of the role of material moves in the working of a language is the key to the rationale of scientific method.²⁸

What Sellars says of science holds a fortiori of any attempt to rationally affirm a philosophical position which claims to be serious about the way the world is.

If the *PNfc*’s need for some rationality norm or other were not important to our argument, what we have said thus far would suffice. However, the argument in chapter six turns on just this point. Therefore, we attempt a general

proof that any rational affirmation of *Nfc* will be conditioned upon some rationality norm.

We have said that rationality norms are necessary conditions for rational affirmations. Rational affirmations are a sub-class of grounded affirmations—they are the grounded affirmations which are not grounded in immediate evidence and/or in truths of logic. Since a proposition which is rationally affirmed is not *directly* derived from what grounds it, a justification—something like a warrant or a license—is needed for affirming the proposition on its more or less remote ground.

If one could bring the grounds for a rational affirmation into immediate relation to it by deduction from evident premises or by reduction to immediate data, then the affirmation would not be a “rational” one in our sense. Instead, it would be one of the stronger sorts of affirmation which philosophers do not argue about. Therefore, without a justification—something like a warrant or a license—no proposition of the sort which philosophers argue about can be affirmed. If the ground could not be linked with the affirmation, the ground would be irrelevant and the affirmation would be groundless.

Of course, it is easy enough to express the link between the ground and the affirmation in logical form. The simplest formal way of doing this would be to make the proposition or propositions which pick out the ground of the affirmation the antecedent of a conditional proposition, and the proposition which is to be affirmed the consequent of the same conditional proposition. Then, since the truth of the ground is given—or assumed—the affirming of the consequent is warranted deductively. However, this formulation of the relationship—or any more complex formulation of it—does not dissolve the difficulty, which merely becomes a question about the truth of the conditional proposition. In many cases, reasons for the truth of the conditional can be given, and the reasoning process can be expressed in further conditionals. But unless the content with which one is dealing is reduced to evidential immediacy or logical equivalence, the question will always remain: Is it reasonable to affirm the proposition in which one was originally interested on its more or less remote ground?

Rationality norms help one to bridge this gap. Whether one prefers with Sellars to call this bridging a “material move” or with Quine a minimal “leap,” or whatever, affirmation under such conditions is an act which can be done reasonably or unreasonably. Rationality norms direct one to a better act of affirming. Rationality norms attempt to say what sorts of characteristics are—to adopt some apt language from Chisholm—“ ‘evidence-making characteristics,’ or ‘reasonability-making characteristics,’ or even ‘epistemically-better-making characteristics.’ ”²⁹

Thus, since *Nfc* is affirmed in the context of *Sfc/Nfc* as a rational affirmation, and since all such affirmations are conditioned by rationality norms, any *PNfc*'s affirmation of his position will be conditioned by such a norm.

One more point about rationality norms is important for our argument. They can be in force or not.

We have made clear that rationality norms are conditions for acts of affirming. In general, conditions for anything can obtain or not, and what they condition obtains only if its necessary conditions obtain. Norms are peculiar conditions; they cannot fail to obtain in the way that empirical conditions can fail to be given. Yet norms can fail to obtain if they are null. Norms can be null, without force, if their being in force entails a state of affairs which itself does not obtain.

Rationality norms, like other norms and rules, regulate. Norms regulate by making a special kind of demand. In the case of many norms, including rationality norms, this demand is not a physical or a psychological exigency. The demand of such norms is that a certain standard be met, although this demand may not be fulfilled. Thus, such norms both set a standard and require that it be met. They not only describe a standard, they prescribe it. But norms cannot actually prescribe anything if the conditions for their fulfillment do not obtain.

Some examples will clarify the notion of actually prescribing—that is, of a norm being in force in contrast with its being null.

One promises his friend, who is in the hospital, to visit on Tuesday. On Monday, the friend dies. One ought to keep one's promises, but in this case the norm is null; that is, it does not actually prescribe. Again, a child is told that he ought to eat his entire doughnut but carefully preserve the hole, since it can be turned in for a free doughnut. Here again, the norm does not actually prescribe; it is null. The child discovers by empirical inquiry that the seeming prescription is incoherent. Again, a professor ought to meet his seminar, but on the way to it his car breaks down. The norm is null; it does not actually prescribe.

In some cases, an apparently valid norm actually is null. An act conditioned upon the norm is performed, and seems to be performed successfully. Yet the act, although done in good faith, is not successful; it is merely putative, and must be set aside. For example, a person is indicted, tried, and convicted of a crime. The charge is based on a section of the criminal code which everyone concerned takes to be in force at the time of the trial. However, the Supreme Court subsequently strikes down this section of the criminal code, by a five to four decision, finding it void because of vagueness. Any person convicted of a crime on a charge laid under this section can obtain a writ setting aside his conviction and ordering his immediate release from prison. His indictment, trial, and conviction were invalid legal acts. They were attempted acts which did not succeed in doing what they were meant to do. They were a merely putative exercise of the processes of justice inasmuch as they were based upon a section of the criminal code which was null and void, a section not in force

—that is, one which did not provide the necessary condition which it seemed at the time to provide for the attempted acts which relied upon it.

In chapter six, we will show what condition must obtain if the rationality norm required for any rational affirmation of Nfc is to be in force. In the present discussion, we think we have clarified two points: 1) that in Sfc/Nfc , Nfc must be rationally affirmed; 2) that the act of rationally affirming Nfc is conditioned upon rationality norms, which must be in force if the attempted act of affirming it is to be successful, not merely putative.