

## 4: Compatibilism

In chapters two and three, we have examined arguments for and against *Sfc*. Neither side has succeeded in establishing its position. This situation naturally raises the question whether the controversy is at all soluble. One possible answer to this question is that the point at issue in the controversy has been misconceived.

We have claimed that *Sfc/Nfc* formulates a central issue in the historical debate about free will and determinism. There are many philosophers, however, who would object that our formulation oversimplifies the debate by allowing no voice to the proponents of the most plausible position: some form of compatibilism. Compatibilists regard their approach as a nuanced solution to the controversy over free will, a solution which avoids confusions or mistaken assumptions common to the *PSfc* and to the extreme—or so-called “hard”—determinist.

In this chapter we discuss the diverse forms of compatibilism and show that they provide no solid ground for formulating the issue otherwise than we do. We argue that soft determinism or reconciliationism—the position often simply referred to as “compatibilism”—either includes *Nfc* or is relevant to the controversy over free will and determinism only to the extent it includes *Nfc*. We also argue against various versions of the thesis that freedom and determinism can be rendered compatible by limiting each to a distinct domain, or by carefully observing the difference between the languages or viewpoints proper to each. We show that all versions of the latter thesis fail on semantical or on formal grounds.

### A. Soft determinism

Soft determinism is the position that free will and determinism are compatible inasmuch as “free” means “uncoerced,” not “uncaused.” This position on the controversy is the one most widely held by English-speaking philosophers. Among those who have articulated and defended it are Hobbes, Locke, Hume, Mill, Moore, Ayer, and Nowell-Smith.<sup>1</sup> Compatibilists in the present century often have defended their position by proposing an analysis of “could have done otherwise” which is consistent with saying it of someone whose act is imputable to him although it is caused.

Most members of this tradition have taken for granted the truth of the scientific worldview and have considered it to involve universal determinism. For example, as Richard Taylor points out, Locke’s whole consideration of the free will issue assumed “that determinism is true and that indeterminism is irrational and unintelligible. The philosophical problem, as he understood it, is simply that of showing that determinism is compatible with what all men believe concerning human liberty.”<sup>2</sup> Even if some soft determinists do not assume the truth of universal determinism, it is clear that they wish to exclude *Sfc*. This is shown by the defenses which philosophers in this tradition make against arguments from immediate experience and moral responsibility for *Sfc*. In chapter two, sections A and B, we described these defenses.

Some reconciliationists reject the characterization of their position as “determinism” or even as “soft determinism.” Perhaps they regard the thesis of universal determinism as false or perhaps they regard it as irrelevant to the initiation of human acts. Nevertheless, reconciliationists hold *Nfc*. If they did not, they would hold nothing which needed to be reconciled with that freedom which they do admit and defend.

It should be noted that when we say that reconciliationists hold *Nfc*, we mean that they reject free choice as we have defined “free choice” in chapter one, section B. Of course, “free choice” as we have defined it signifies a mode of freedom which is rejected by some proponents of free will. For example, anyone who accepts the principle of sufficient reason—as many proponents of free will do—is likely to reject the position that a person’s very choosing is one of its own necessary conditions. Such a proponent of free will might regard himself as a soft determinist of the sort classically described by William James—that is, as one who holds both that determinism obtains and that moral responsibility is real.

Apart from their rejection of *Sfc*, compatibilists of this sort seem to us to affirm what is true. There are senses of “freedom,” “responsibility,” “imputability,” “voluntariness,” and so on, compatible with *Nfc*. Moreover, these words often are used in such senses, especially in legal and social contexts. It is

not clear that law requires any mode of imputability other than that which a sophisticated compatibilist can admit. Aristotle's concept of voluntariness, for example, seems both consistent with *Nfc* and adequate for legal purposes.

Thus, in our view, "freedom" does have a sense compatible with *Nfc*. In fact, as we make clear in chapter one, section A, "freedom" has many such senses, for it can signify physical freedom, freedom to do as one pleases, ideal freedom, political freedom, and creative freedom.<sup>3</sup> But in chapter one, section B, we also define another mode of freedom: free choice. Clearly, no compatibilism which includes *Nfc* can allow that "freedom" used in this sense refers to anything real. Yet "freedom" can be used in this sense, even if mistakenly, to refer to a supposed property of choice, the data of which we also describe in chapter one, sections C through F. As we show in chapter two, section A, these data can be interpreted in a way consistent with *Nfc*, but as arguments for free choice based upon immediate experience make clear, these data also can be interpreted in a way inconsistent with *Nfc*.<sup>4</sup>

The fact that the data of choice can be interpreted in a way consistent either with *Sfc* or with *Nfc* establishes *Sfc/Nfc* as a genuine controversy. The issue in this controversy cannot be evaded by pointing to other senses of "freedom" in which the assertion that one is free is consistent with *Nfc*, any more than the demand for political freedom can be evaded by pointing out that oppressed people can be called "free" in other, irrelevant senses.

If someone straightforwardly asserts soft determinism as a position which excludes *Sfc*, then, no matter what other modes of freedom he asserts to be real, he asserts *Nfc*, that is, he denies to be real the very freedom at issue in *Sfc/Nfc*. Therefore, if this is a soft determinist's position, the compatibility he defends is not relevant here; his position does not state that *Sfc* and *Nfc* are compatible.

As the exposition in chapter two, sections A and B, makes clear, an important element of the soft determinist's position, as it developed historically, is the contention that "freedom" means *only* physical freedom and freedom to do as one pleases. Hume's statement of the position is classic. Anyone not a prisoner and in chains is free. For Hume and those who follow him, "free" does not mean anything more than "not compelled" or "not coerced."

Despite its plausibility to many philosophers even today, this line of argument is fallacious. It is not fallacious in indicating that physical freedom and freedom to do as one pleases are compatible with *Nfc*; they are. It is not fallacious simply because it implies that no one can make a free choice; this is a position for which one can argue. But it is fallacious in arguing that the controversy over free choice is dissolved by pointing out the compatibility of determinism with physical freedom and freedom to do as one pleases. There remains another *possible* sort of freedom—free choice.

"Freedom," however, is not the only relevant expression which has several

distinct meanings. “Responsibility” likewise has a number of meanings, and the various senses of “responsibility” correspond to various senses of “freedom.” Language used in evaluative statements also has multiple senses, and in at least some cases these senses correspond to distinct senses of “free” and “responsible.” If one uses any one of these expressions in a given context in a certain sense, he is likely to use the related expressions in a corresponding and appropriate way. Thus, if a three-year-old child has a tantrum and purposely breaks something, one might say that the child is “acting freely,” is “responsible,” is “naughty,” and deserves “punishment.” One might use the same expressions in referring to a president who obstructs justice, except that one probably would say “did wrong” rather than “was naughty.” Many people using these expressions in such different contexts would use them with different senses, believing that the president was free and responsible in a sense in which a three-year-old child could not be. In each context, all of the relevant expressions would systematically shift in meaning.

Therefore, since there is a *possible* mode of freedom—free choice—in addition to those for which the soft determinist can account, there will be corresponding senses of the other relevant expressions for which he cannot account. Thus, although the soft determinist shows that “free,” “responsible,” and related expressions can be correctly used in talking about human action in senses compatible with *Nfc*, he does not succeed in dissolving *Sfc/Nfc*. *Sfc* remains incompatible with the theory of universal determinism and with any other theory which excludes free choice as impossible.

One of the common arguments in favor of *Sfc* has been that a person cannot be morally responsible for his acts unless it is possible for him either to conform to a norm or to violate it. This is the meaning of the assertion: “ ‘Ought’ implies ‘can.’ ” To defend their position against this argument, compatibilists in recent years have tried to provide an adequate analysis of “I could have done otherwise” compatible with *Nfc*. Candidates for such an analysis include “I would have done otherwise had I so chosen,” or “had I tried,” or “had circumstances been slightly different.” These are possible meanings of “I could have done otherwise.” They express what people sometimes mean when they utter these words. These expressions can be used to refer to physical freedom, to freedom to do as one pleases, or to both.

But these are not the only possible meanings of “I could have done otherwise.” “Could” also can correspond to the “can” in “I can choose either *A* or *B*.” As we make clear in chapter one, section G, “can” might be used by a person to express the belief—true or false—that there are alternatives to be settled by his choice alone. In other words, the description of the experience of choice in chapter one, sections C through F, exhibits a foundation for a meaningful use of “can” incompatible with *Nfc*.

This particular use of “can” cannot be analyzed in a way inconsistent with

*Sfc*. To refuse to admit that “can” might sometimes be correctly used in a way inconsistent with *Nfc*—that is, to affirm *Sfc*, even if this affirmation is false—is to maintain that stipulation can resolve a substantive question.<sup>5</sup>

However, because of the impressive array of philosophers who have espoused soft determinism, it seems unlikely that there is nothing more in the case for this position than a stipulation that expressions such as “freedom” and “responsibility” be defined in a way consistent with *Nfc*. Thus, it is worth asking why a soft determinist would suppose that he has disposed of the sense of “freedom” required to affirm *Sfc* merely by showing that “free” and certain related expressions have at least one meaning compatible with *Nfc*.

One possible answer to this question is suggested by our discussion in chapter two, sections A and B, of the arguments for *Sfc* from immediate experience and from moral responsibility. The argument from immediate experience fails because there are senses of “freedom” compatible with the phenomena of choice and with the ordinary language required to describe these phenomena, as well as with the denial of *Sfc*. The argument from moral responsibility fails because there is a sense of “moral responsibility” compatible with *Nfc*. In criticizing these arguments, a *PNfc* need only show that there are in fact distinct senses of “free” and “responsible.” He need not show that no one ever uses these expressions in propositions inconsistent with *Nfc*.

The experience of choice does not by itself demonstrate the reality of free choice, since—as we have made clear—an interpretation of the phenomena of choice compatible with *Nfc* remains a logical possibility. To point to this logical possibility is sufficient to show the invalidity of the argument: “Since one makes choices, one makes free choices.” Likewise, moral responsibility is not conclusive evidence for *Sfc*, because there is a plausible *sense* of “responsible” compatible with *Nfc*.

Nevertheless, the soft determinist calls attention to the possible use of “free,” “can,” and so on, in a sense compatible with *Nfc* in contexts other than that of answering unsound arguments for *Sfc*. The soft determinist characteristically argues as if to note meanings of the relevant expressions compatible with *Nfc* were sufficient to rule out the notion of free choice as in some way confused or suspect. But, clearly, more must be involved in such arguments than the mere observation that these words sometimes have meanings compatible with *Nfc*.

To make clear what might be involved in such a claim, we note that all the key terms in the discussion have parallel shifts in their meanings. If a soft determinist ignores this point, he might suppose, for example, that an account of moral responsibility which involves *Nfc* renders superfluous the sense of “responsibility” which includes “free choice” in its definition. Moreover, assuming that “responsibility” has only one meaning, the soft determinist might think that given his deterministic account of moral responsibility, an

analysis of “responsibility” involving “free choice” is necessarily mistaken.

In fact, however, the soft determinist accounts not for every *possible* mode of responsibility, but only for a certain mode of it. As we explained above, many people predicate “morally responsible” of a naughty three-year-old child and of a felonious president in different senses, denying that the former made a free choice and claiming that the latter did so. Whether or not such a claim can be correct, the meaning of “morally responsible” which includes “free choice” is given in the very making of the claim.

A soft determinist might grant that the child and the president are called “morally responsible” in different senses, but maintain that both senses are susceptible to an analysis consistent with *Nfc*. But some *PSfc* might claim that in asserting moral responsibility of the president he uses “moral responsibility” in a sense which includes “free choice” in its definition. Surely, such a claim is *intelligible* even if it is in principle false because *Nfc* is true. One who makes this claim wishes to insist that the president could have avoided obstructing justice but freely chose to obstruct it—“freely chose” as defined in chapter one, section B.

A soft determinist might point out that inasmuch as he has been able to provide analyses of many senses of “can” which are compatible with *Nfc*, further analysis might show that even the sense of “can” involved in the definition of “free choice” is compatible with *Nfc*.<sup>6</sup> This suggestion, however, depends on the view—which is a curious view of the relationship between the meaning which speakers intend when they use language and the meaning which analysis can uncover—that analysis can show that words really mean the contradictory of what speakers in using them mean to say.

Of course, analysis can show that a speaker’s expressions convey meanings he did not intend. What a person says often has implications to which he fails to attend. If such implications are pointed out, a person often withdraws his original statement and substitutes for it one which avoids the unintended implications. But analysis cannot show that a speaker’s expression has a “true meaning” which is logically incompatible with the meaning he consciously intends to express by using the precise language he does.

In the preceding paragraphs, we have been trying to understand why the soft determinist assumes that his articulation of meanings of “free”—and some related expressions—compatible with *Nfc* rules out the senses of these expressions which involve free choice. The preceding clarification reveals the inadequacy of some possible reasons for this assumption. But the inadequacy is so glaring that there must be a different reason why most soft determinists make this assumption.

In fact, there is a different reason: Soft determinists believe *Nfc* true. Some of them, like Hobbes, accept *Nfc* on metaphysical grounds; many others, like Locke and Hume, accept it as a deliverance of science. Considering the

historical context in which soft determinism has flourished, one can understand how universal determinism has seemed an obvious truth admitted by all men. Thus, as we have noted at the outset of this section, determinism is generally assumed by soft determinists.

Even more basic than the influence of its historical context, however, is the context of controversy in which the reconciliationist project has its place. Reconciliation is called for only if seemingly incompatible data are given. If *Nfc* were not assumed to be true, there would be no need to reconcile human freedom with something else. Nothing would oppose *Sfc*. Thus, soft determinism is not an initial position in the controversy over free will, but is rather a countermove, an attempt to meet objections to an initial position which includes *Nfc*.<sup>7</sup>

The assumption of *Nfc* by the soft determinist also is revealed by the fact that a soft determinist frequently makes statements which are commonly made by a *PNfc* in arguing for his position. For example, soft determinists say that there can be no doubt that every choice is caused, that a free choice would be a random and unintelligible event, that responsibility requires that one's choices be determined by one's character, and so on.

Clearly, the soft determinist's assumption of the truth of *Nfc* provides no ground for excluding the meaning of "free" required to affirm *Sfc*. If one of the contradictory propositions is true, the other, even if false, is meaningful. The truth of *Nfc* would provide no ground for regarding the use of expressions needed to state its contradictory as suspect or confused.

Nothing in our discussion of soft determinism presupposes the truth of *Sfc*. Our point merely is that "free choice" is meaningful and that *Sfc/Nfc* cannot be dissolved by talking about kinds of freedom the admission of which is compatible with *Nfc*.

## B. The double-aspect theory

Soft determinism is not the only attempt to show that *Sfc/Nfc* is a misformulation of the controversy over free will. There is another form of compatibilism which has been adopted by many philosophers since Kant. We call this form of compatibilism "the double-aspect theory."<sup>8</sup>

On this view, determinism does not involve *Nfc*, because the language appropriate to determinism cannot be used to refer to choice. One can refer to choice only from the standpoint of the agent or in the practical language of action. Determinism is a thesis about events; it can make reference only from the standpoint of an observer or in a theoretical language such as that of science. The two distinct aspects have been called "distinct levels of description and explanation," "diverse language strata," "different bodies of discourse," "separate domains," and so on.

A contemporary example of the double-aspect theory is found in A. I. Melden's *Free Action*:

Where we are concerned with causal explanations, with events of which the happenings in question are effects in accordance with some law of causality, to that extent we are not concerned with human actions at all but, at best, with bodily movements or happenings; and where we are concerned with explanations of human action, there causal factors and causal laws in the sense in which, for example, these terms are employed in the biological sciences are wholly irrelevant to the understanding we seek. The reason is simple, namely, the radically different logical characteristics of the two bodies of discourse we employ in these distinct cases—the different concepts which are applicable to these different orders of inquiry.<sup>9</sup>

Melden's statement of this theory is typical of expressions of it by many other authors.<sup>10</sup>

It is possible to hold the double-aspect theory for reasons irrelevant to the controversy over free choice. One might hold that there is nothing about either domain incompatible with *Sfc*. One might also hold that both domains exclude *Sfc*. The latter view might take the form of maintaining causal determinism for events and determination by reasons for actions. Such a view clearly would include *Nfc*. Thus, it is not a form of compatibilism. In the remainder of this section, we are concerned only with proponents of a double-aspect theory who think that the distinction of two domains permits them to assert the compatibility of freedom with determinism, and thus to avoid the contradictory alternatives expressed in our formulation of the controversy.

Construed in this way, the double-aspect theory still requires clarification. What is the relationship between the two domains? Is the distinction between the two bodies of discourse merely a matter of fact? That is, can the distinction be eliminated by a more adequate understanding of the rules by which each body of discourse might be translated into the other? If so, the double-aspect theory does not dissolve *Sfc/Nfc*. The *PNfc* will be able to look forward to the discovery of the required translation rules so that he will be able to use theoretical language to refer to choices and to affirm *Nfc*.<sup>11</sup> The *PSfc*, for his part, can admit that theoretical language might be used to refer to choice, but he will use such language to deny the truth of *Nfc* and whatever other theoretical propositions he must deny to maintain *Sfc*.

On another interpretation, the double-aspect theory claims that the two bodies of discourse are in principle irreducible to one. On this interpretation, each language is indispensable for certain purposes and neither language can be translated into the other. Such a view can take either of two forms.

In one form, it might involve the claim that reference, truth, and explanation are possible in only one body of discourse—that appropriate to talking about events theoretically and in terms of factors which exclude *Sfc*.<sup>12</sup> On this

approach, the body of discourse appropriate to action contains no statements and thus allows neither the *PSfc* nor the *PNfc* to affirm anything other than what can be expressed in theoretical language. It follows that this interpretation of the double-aspect theory does not permit the claim that freedom and determinism are compatible by virtue of the distinction between bodies of discourse.

In another form, the double-aspect theory which holds that there are irreducible bodies of discourse allows truth-claims to be made in both of them. Only on this approach can the double-aspect theory permit both freedom and determinism to be truly affirmed, yet avoid the seeming contradiction between them by positing the irreducibility of the domains proper to each. But even in this form, the double-aspect theory will not put to rest *Sfc/Nfc*.

The *PNfc* can admit the irreducibility of the language of action to that of causally determined events. Still, he might argue, the former language is dispensable and there are good reasons to dispense with it, at least to the extent that it is not translatable into deterministic language.<sup>13</sup> Such reasons might include the explanatory purposes of science, the promotion of reform of criminal law, and the practical confusion arising from the use of the language of action. For his part, the *PSfc* must of course resist any attempt to eliminate the language of choice.

Thus, this form of the double-aspect theory fails to dissolve *Sfc/Nfc*. It fails because the analysis of the use of ordinary language cannot by itself settle a substantive philosophical question. Of course, such analysis is essential for the accurate formulation of the issue. We grant that it is significant that ordinary language about human actions articulates the phenomena of choice in such a way that this language is not altogether reducible to scientific discourse. It does not follow, however, that the *PNfc* is necessarily misguided or wrong-headed. The most which follows is that if *Nfc* is correct, people should more or less radically revise the way they talk about human actions.

Some advocates of the double-aspect theory would point out at this juncture that their view is not merely descriptive of the way language does work. They would claim that the irreducibly distinct bodies of discourse regarding actions and regarding events are indispensable because without them and their distinction one cannot speak coherently about human behavior. Such a claim expresses in terms of language the same thesis which Kant expresses by speaking of the irreducibility of theoretical and practical thinking.<sup>14</sup> The conditions of the possibility of meaningful discourse and of coherent thinking either are the same or—at least in respect to this problem—precisely analogous. It is no accident that many proponents of the double-aspect theory remind one of Kant even if they do not mention him.

Because Kant's treatment of this form of the double-aspect theory is clear and well-developed, we first consider the proposal in his own terms.

In the *Critique of Pure Reason* (A444-445=B472-473), the thesis of the

third antinomy is: "Causality in accordance with laws of nature is not the only causality from which the appearances of the world can one and all be derived. To explain these appearances it is necessary to assume that there is also another causality, that of freedom." The antithesis is: "There is no freedom; everything in the world takes place solely in accordance with laws of nature."<sup>15</sup>

The arguments Kant proposes for the thesis and the antithesis are not important for our present purpose. What is important is the solution he offers to the antinomy. Because of its difficulty, this solution has been interpreted in various ways. We offer the following as a plausible interpretation, relevant to the double-aspect theory.

Kant's solution to the antinomy begins (A532=B560) with the assertion that there are two and only two kinds of causality conceivable by the human mind: causality according to nature and causality according to freedom. Causality according to nature presupposes time, and it leads to a causal chain or network covering all the facts of the natural world. Causality according to freedom is the power of initiating something spontaneously.

Having made this distinction, Kant proceeds to argue that if all causality in the sensible world were mere nature, then there would be no room for practical freedom, which presupposes that something can happen which *ought* not to have happened. He attributes to man a power of free choice which can be affected by sensuous impulses but which cannot be coerced by them (A533-534=B561-562).

The key to a solution, according to Kant, is not that there is any gap in the solid network of natural causality. Rather, he thinks, the key is that the whole of nature is a world of appearance, not of absolute reality. If nature is not ultimate reality, something nonnatural can be the term of a relationship of a natural event. Thus, Kant is able to formulate the problem in the following terms: "Is it a truly disjunctive proposition to say that every effect in the world must arise *either* from nature *or* from freedom; or must we not rather say that in one and the same event, in different relations, both can be found?" (A536=B564).

Kant next argues that the causality of a person can be regarded from two points of view. To the extent that the person is an appearance in the sensible world, its causality can be regarded from one point of view. To the extent that the person also has a faculty which is not an object of sensible intuition, but through which it can cause appearances, its causality can be regarded from another point of view. Considered as a causality of a thing in itself, this causality is *intelligible* in its action. But considered as a causality belonging to an appearance in the natural world, this causality is *sensible* in its effects.

Thus, for Kant, the human person has an empirical aspect, and according to this aspect the person's causality and effects in the world of nature are altogether determined by antecedent causes. At the same time, this same person has an intelligible aspect, and according to this aspect nothing happens to the

person, no action begins in the person, and there can be no change in the person, “but we may yet quite correctly say that the active being *of itself* begins its effects in the sensible world. In so doing, we should not be asserting that the effects in the sensible world can begin of themselves; they are always predetermined through antecedent empirical conditions . . .” (A541=B569).

Might it not be the case that in certain instances the empirical causality is an effect of an intelligible causality (A544=B572)? Of course, one would have to assume that “the *action* of these causes *in the appearance* is in conformity with all the laws of empirical causality” (A545=B573). Man knows himself in respect of intelligence and reason to be a purely intelligible entity (A547=B575). The fact that reason has causality is evident from imperatives which persons impose on their active powers; nature knows only “is,” “was,” and “will be,” not “ought” (A547=B575).

Having reached this point, Kant’s attempted explanation breaks down. “Ought” expresses a conceptual ground of action; a merely natural action always has an empirical causal condition. “The action to which the ‘*ought*’ applies must indeed be possible under natural conditions,” Kant affirms. But the conditions do not determine the will. Whatever is willed, whether the pleasant or the good, “reason will not give way to any ground which is empirically given,” for it follows its own order of ideas “to which it adapts the empirical conditions.” If sensuous impulses impel me to will, they cannot give rise to an “ought”; reason declares actions to be necessary, “although they have never taken place, and perhaps never will take place” (A548=B576).

Here Kant is discussing two different things. On the one hand, reason may issue its own demands, whether these be fulfilled or not. Reason says what ought to be, and this necessity, Kant points out, is irreducible to natural requirements and conditions. On the other hand, actions which are presumably caused by reason occur in the empirical world. This world must be such as to allow for their possibility; the empirical conditions must be adaptable to the intelligible causality. This discussion leaves unclear how empirical conditions which are woven into the unbroken fabric of the totality of nature can be adapted for particular acts.

Reason itself, Kant goes on, must have an empirical aspect. This can be completely investigated from experienced actions, which reveal the subjective principles of an individual’s will. But if we compare the acts—which from an empirical point of view are inevitable—to reason as a cause, Kant says that perhaps what inevitably happened ought not to have happened. Then Kant makes the remarkable statement: “Sometimes, however, we find, or at least believe that we find, that the ideas of reason have in actual fact proved their causality in respect of the actions of men, as appearances; and that these actions have taken place, not because they were determined by empirical causes, but because they were determined by grounds of reason” (A550=B578).

The problem with this statement is not whether it is true, but whether it can make sense at all on Kant's own principles. On the one hand, the only actions which can be known are effects in the world of experience; everything in that world is not *more or less* determined, but *completely* determined. Yet Kant is saying that some actions have taken place "not because they were determined by empirical causes." On the other hand, "ought" points to an intelligible causality, and nothing happens to this principle in itself—it is an unchanging principle. But actions which take place in the empirical world—a world uniformly determined as a completely interlocking system—sometimes are determined by grounds of reason and sometimes are not. Kant here breaks down his own distinction between the two domains; he treats the causality of nature and the causality of the noumenal self as if they were each necessary conditions but only together a sufficient condition for those empirical effects which are human actions.

A confirmation of the correctness of our criticism is found in a footnote in which Kant argues that since a person does not know the intelligible character except by its empirical effects, he can never know the real morality of actions, even of his own. "Our imputations can refer only to the empirical character. How much of this character is ascribable to the pure effect of freedom, how much to mere nature, that is, to faults of temperament for which there is no responsibility, or to its happy constitution (*merito fortunae*), can never be determined . . ." (A551=B579). Kant should keep the two modes of causality distinct; in this footnote, he mingles them.

In the text, Kant goes on to argue that the causality of reason is outside time; it therefore can be a spontaneous starting point of a new series of effects. The same cause—man, the moral agent—belongs to nature, and from this point of view "no given action (since it can be perceived only as appearance) can begin absolutely of itself" (A553=B581). To say that the act as part of nature cannot begin absolutely of itself is an understatement on Kant's principles; he should say that the act does not at all begin of itself, but is wholly determined by antecedent conditions.

Kant illustrates his point with the example of a voluntary action, a malicious lie. He says that one can trace the empirical aspect of the action to its sources, but "we none the less blame the agent" disregarding conditions "just as if the agent in and by himself began in this action an entirely new series of consequences." One proceeds on a law of reason according to which the agent can and ought to have determined, irrespective of all the empirical conditions, to act otherwise.

This causality of reason we do not regard as only a co-operating agency, but as complete in itself, even when the sensuous impulses do not favour but are directly opposed to it; the action is ascribed to the agent's intelligible character; in the moment when he utters the lie, the guilt is entirely his. Reason, irrespective of all

empirical conditions of the act, is completely free, and the lie is entirely due to its default. (A555=B583)

Here, the moral judgment, the possibility of which is denied four pages previously, is firmly made. Moreover, whereas four pages previously Kant says that imputations can be made only to the empirical character, here the action is ascribed to the agent's intelligible character.

It might be thought that Kant's difficulties arise only because of certain distinctions not yet clearly made in the first *Critique*, distinctions which Kant made later on. For our present purpose, it is not essential that the analysis of Kant's works be complete. To answer the objection, it is sufficient to quote the following from Lewis White Beck, a leading scholar sympathetic to Kant:

If by "freedom" we mean noumenal causation and assert that we know no noumena, then there is no justifiable way, in the study of phenomena, to decide that it is permissible in application to some but not others of them to use the concept of freedom. The uniformity of human actions is, in principle, as great as that of the solar system; there is no reason to regard statements about the freedom of the former as having any empirical consequences. If the possession of noumenal freedom makes a difference to the uniformity of nature, then there is no uniformity; if it does not, to call it "freedom" is a vain pretension.<sup>16</sup>

In his effort to dissolve *Sfc/Nfc*, Kant assigns determinism to the natural sequence of causes and effects, while he postulates free choice in the principle of action of a moral subject. This separation would be sufficient to solve the problem only if the two domains were kept completely distinct. The danger is in attempting to use the two viewpoints simultaneously in order to provide something like a stereoscopic view of the reality of human action.

But Kant precisely makes such an attempt. He wants to be able to look at human actions which occur within experience and to consider them at the same time as effects of a causality outside the unbreakable network of natural causes and effects.

The result is that sometimes Kant thinks of action precisely as an appearance; he then relates it to other appearances, and announces that no one can ever know the moral quality of anyone's action, including his own. Sometimes, however, Kant thinks of action as an expression of moral agency, and then he is willing to say that the malicious liar is guilty regardless of the whole sequence of conditions which made his act—as a piece of empirical behavior—as inevitable as anything in nature can be. Kant also sometimes tries to relate an action simultaneously to empirical causes and to moral agency, regarding the action as somehow composite and as somehow an effect of both sorts of principles—for example, when he speaks of mitigated responsibility.

When Kant is consistently looking at action from a single point of view, it is difficult to see how he could establish any empirical criteria according to which

any observable sequence of behavior would count as a “malicious lie,” or, in general, what criteria Kant could offer for distinguishing any set of phenomena and regarding it as an act to be imputed to a moral agent. When Kant does not consistently look at action from a single point of view, it is difficult to see how he overlooks the inconsistencies into which he falls.

Kant could avoid these inconsistencies and maintain a place for both *Sfc* and *Nfc* only if he firmly refused to allow anything whatsoever to fall into both domains. Yet he wishes to use his distinction to deal with human action and he does not wish to restrict action to one of the two domains.

One reason why Kant’s version of the double-aspect theory is interesting is that in itself it is a theory not of two bodies of discourse, but rather of two standpoints for thinking. In reading Kant, one easily notices the distinction between ordinary language and the conditions for the possibility of meaningful discourse. Kant continually uses ordinary language in a way which violates the requirements of his double-aspect theory. We think that recent proponents of the double-aspect theory make a similar mistake, but their lack of explicitness in stating their position conceals the mistake, for it is hard to tell whether they are maintaining a thesis about ordinary language or about what is necessary for any meaningful discourse about human behavior.

A proponent of the double-aspect theory might answer that the inconsistencies we have pointed out in Kant’s text are peculiar to Kant and reveal no underlying problem with the double-aspect theory. We admit that Kant’s approach has its idiosyncracies. But we think that Kant’s inconsistencies are a symptom of a problem which confronts the double-aspect theory as such.

If the incompatible predicates, “free” and “determined,” can be said of the same human behavior or anything necessarily related to the same behavior, then there will be something of which it is possible to say that it is both free and determined. Thus, even if Kant’s inconsistencies are avoided, similar inconsistencies will emerge—or, at least, the double-aspect theory will not preclude their emergence. To preclude such contradictions, the bodies of discourse must be separated to such an extent that nothing whatever can be the common referent of the relevant expressions in the distinct languages. Thus, the theory requires what J. R. Lucas calls “a thorough-going schizoglossia which is blatantly at variance with the facts.”<sup>17</sup>

But if the two bodies of discourse are separated to this extent, it will not be possible to refer to anything which is a piece of human behavior as the referent or as a component of the referent of “free” and “determined.” In this case, the double-aspect theory is pointless. It is pointless because the apparent contradiction which the double-aspect theory was intended to remove could never arise if it were not possible—at least mistakenly—to refer to the same thing as both “free” and “determined.”

But the proponent of the double-aspect theory might object that the whole

point of his approach is to eliminate confusion by showing that expressions of the two bodies of discourse do not have the common reference often mistakenly assumed for them. This objection involves a very strange view: that without making any false affirmations one can use expressions to make affirmations and be mistaken about their reference by supposing that two expressions have a common reference which they do not have. The double-aspect theorist is trying to tell a *PNfc*, not that he is in error in affirming that actions are determined, but that when he says “actions” he is really referring to something other than actions.

The implausibility of the double-aspect theorist’s view can be made clear by an example. Imagine a prosecutor outlining his case at the opening of a murder trial: “Discovering that he would inherit his uncle’s estate, the defendant deliberately planned to kill the victim, his uncle. He took a gun, held the gun to his uncle’s head, and pulled the trigger. The victim’s brains spilled out, and he died instantly.” The double-aspect theorist must defend at least the following proposition: It involves a category mistake to say, even falsely, that in the circumstances precisely as they were, the trigger moved *because* the nephew deliberately planned to kill his uncle.

A philosopher might try to maintain this proposition, but to do so would be to deny the very possibility of any necessary relationship between the victim’s death (a physical event) and the murderer’s deliberate plan—part of a human act. In short, the trouble with the double-aspect theory is that even if human acts are not precisely physical events, there is a necessary relationship in given circumstances between certain physical events and a certain human act. The logic of the double-aspect theory demands that its proponents say that the uncle’s death and the nephew’s deliberate plan had nothing to do with each other.<sup>18</sup>

Thus, Kant’s difficulty was not due to any idiosyncrasy. Whether one uses Kant’s phenomenon/noumenon distinction, or whether one uses a reasons-language/causes-language distinction, or some other distinction, one nevertheless has the same basic difficulty as Kant, provided that the two domains are defined by their opposition and are distinguished precisely in order to avoid the contradiction which would otherwise arise.

Of course, as we said in the earlier part of this section, someone might use such a pair of expressions to make a distinction not meeting the specified conditions. In particular, he might distinguish reasons-language from causes-language without assigning determinism to the latter and some principle incompatible with determinism to the former. But in this case even the appearance of compatibility between determinism and freedom is lost. Our formulation of *Sfc/Nfc* is not shown to be mistaken.

A double-aspect theorist might object at this point that our entire discussion

of the double-aspect theory thus far has been based upon a misunderstanding of it. He might claim that he can slip between the horns of the dilemma we have posed for him. The common reference he needs is achieved while the contradiction we point out is avoided, if the same thing can be considered in two bodies of discourse from different points of view. He believes this possible and holds that in one perspective a piece of human behavior can be seen as a determined event while in another perspective the same thing can be seen as a free act. The behavior seen as the trigger being pulled in the event-perspective is seen in the action-perspective as the pulling of the trigger. The difference in perspective removes the contradiction, without precluding common reference.<sup>19</sup>

But this approach does not remove the contradiction; it only appears to do so. The appearance is created by the metaphor of “different points of view.” The notion of points of view is based on an analogy between vision and propositional knowledge.

Different visual points of view on the same object present, so to speak, different pictures of one and the same thing. These pictures can be radically different—for example, the view of a coin on edge and the view of the coin’s face. Yet these pictures do not contradict one another; they are simply different. They do not conflict precisely because they are *pictures* of the object; they make no claims about the object itself.

Propositional points of view are like visual points of view in some ways, but are different in a crucial respect: The propositions which collectively constitute a point of view on some subject matter are not pictures of the subject matter. If they are affirmed, they are *claims* about the subject matter itself. It is one thing to say that a coin on edge looks like a two-dimensional rectangle; it is another thing to affirm that the coin on edge *is* a two-dimensional rectangle.

What perhaps lends plausibility to the analogy is the way in which certain propositions are expressed. For example, “He was standing to the right of the desk” and “He was standing to the left of the desk” seem to be contradictory, but both could be true, if stated by persons who viewed the situation from different visual points of view. But this use of the expression “different points of view” is simply another way of expressing the requirement of the principle of noncontradiction usually expressed by “in the same respect.”

Thus these two statements, if both true, are *not* expressions of contradictory propositions. Either they are not fully explicit statements of the same proposition or they are not fully explicit statements of different propositions. If they are statements of the same proposition, then the difference in point of view has been discounted and no contradiction is involved; if they are statements of different propositions, then the points of view are included in the propositions as part of the states of affairs being described. In the latter case, since the subjects of predication are different, the propositions cannot be contradictory.

Thus, the difference of points of view construed in this way is of no use to the double-aspect theorist; *Sfc* and *Nfc* are contradictories; they refer to the same thing in the same respect.

The flaw in the points-of-view analogy can now be explicated. Different visual points of view produce different pictures of the same thing. By treating propositional knowledge as if it were vision, one easily takes for granted that contradictory propositions are merely different pictures of the same thing. On this analogy, contradictory propositions would be incompatible only if they were affirmed from the same point of view. If they are affirmed from different points of view, however, such propositions are only different pictures of the same thing.

But here the analogy is carried too far. A proposition affirmed from a given propositional point of view about something other than the point of view itself does not characterize the point of view. Instead, someone uses it to pick out some other state of affairs and affirms that this state of affairs obtains—obtains independently of the conditions of one's knowing it and talking about it. Thus, the contradictory propositions affirmed about the same state of affairs from different propositional points of view are no less contradictory for their being affirmed from these different points of view. These propositions are not about the points of view from which they are affirmed, but about some other state of affairs; what they articulate is that state of affairs, independent of anyone's knowing it and talking about it. If it happens that the point of view of one affirming a proposition is confused by him with what he is talking about, then he makes false statements about the world.<sup>20</sup>

Our point can be restated more briefly. The difference between visual points of view depends on the conditions for seeing, not on what is seen. Propositional knowledge about something, however, claims to articulate states of affairs, not the conditions of one's knowing them to be so. Different propositional points of view are precisely different conditions for one's knowing and talking about things. Hence the difference of points of view in this case makes no difference at all; differences in propositional points of view are precisely excluded by the claim involved in the affirming of any proposition: that the state of affairs which it picks out obtains.

The application of this analysis to the example of the two expressions, "the pulling of the trigger" and "the trigger's being pulled," makes clear that they do not belong to different bodies of discourse whose distinction allows statements which would otherwise be contradictory to be affirmed without contradiction. The two expressions can be used in stating the same proposition; if so, only a single statement is made. This difference of point of view is merely grammatical. The two expressions can be used in stating different propositions; if so, they either refer to the same thing or not.

If to the same thing, the two propositions will contradict each other provided

that the properties which are attributed are incompatible. Thus, if “The pulling of the trigger was a free act” and “The trigger’s being pulled was a determined event” refer to the same thing, then the conjunctive proposition expressed by these two sentences will be logically impossible, assuming that “free act” and “determined event” signify incompatible properties, as they do if something’s being a determined event entails that, whatever else it might be, it is not a free act.

If the expressions are used to refer to different things, then the two propositions could not be contradictory and the distinction between viewpoints is unnecessary. If “the pulling of the trigger” refers to a human act and “the trigger’s being pulled” refers to an object’s movement, then to say that the one is free and that the other is determined does not even seem to be a contradiction.

In sum. The double-aspect theorist cannot escape the dilemma we have pointed out. Either he fails to avoid the contradiction between *Sfc* and *Nfc*, or he avoids contradiction at the price of making his distinction pointless. The double-aspect theory, however subtly developed, is no more successful than soft determinism in its attempt to show the dissolubility of the controversy over free will. Soft determinism includes *Nfc*, while the double-aspect theory fails to preclude it.

In chapter two, we showed that the existing ways of arguing for *Sfc* are unsuccessful. In chapter three, we showed that the case for *Nfc* is a weak one. But we are not skeptics. We think that *Sfc/Nfc* can be resolved in favor of *Sfc*. In chapter five, we articulate the method to be used in refuting *Nfc*. This refutation will provide the key to establishing *Sfc* in chapter six.