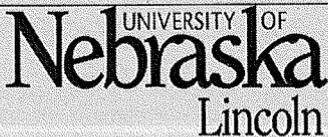


McLear, Colin (cmclear2)



Request an Article

ILLiad TN: 832313



Call # B2798 .S39 1984

Journal Title: **Self and Nature in Kant's Philosophy**

Volume: Issue:

Month/Year: **1984**

Pages: **73-101**

Article Author: Allen Wood

Article Title: *Kant's Compatibilism*

Location: LOVE

4/29/2014 7:20 AM

Photocopied materials are all delivered electronically.
Those items that are not of adequate quality for scanning
will be mailed at library discretion.

Sent

Updated

III

Kant's Compatibilism

Allen W. Wood

On the issue of free will and determinism, philosophers are usually categorized either as "compatibilists" or as "incompatibilists." Compatibilists hold that our actions may be determined by natural causes and yet also be free in the sense necessary for moral agency and responsibility. Freedom and determinism are compatible. Incompatibilists hold that if our actions are determined to take place by natural causes, then free agency and moral responsibility are illusions. Freedom and determinism are incompatible.

Kant's views on many issues do not fit neatly into the customary pigeonholes, and the free will issue is no exception. He is probably most often regarded as an incompatibilist, and not without justification. Certainly he repudiates Leibniz's compatibilism, which (Kant says) allows us no more than "the freedom of a turnspit" (*KpV* 97g 101e).¹ Yet Kant himself holds that not only are our actions determined by the causal mechanism of nature and but also that they are free. The basic question, he says, is "whether regarding the same effect which is determined by nature, freedom can nevertheless be present or whether freedom is wholly excluded by such an exceptionable rule" (*A536/B564*). In answering this question in favor of the former alternative, Kant's avowed purpose is "to unite nature and freedom," to

¹All translations are my own.

"remove the apparent contradiction between the mechanism of nature and freedom," to show that "causality from freedom at least does not contradict nature" (*A537/B565; KpV* 97g 101e; *A558/B587*). But these are precisely the aims of compatibilism. When we consider all Kant's views together, it is tempting to say that he wants to show not only the compatibility of freedom and determinism, but also the compatibility of compatibilism and incompatibilism.

In brief, Kant's theory is that our actions may be simultaneously free and causally determined because we belong to two worlds; this means that our existence can be viewed from two different standpoints. We are, on the one hand, part of nature, of what Kant calls the sensible or phenomenal world. We are objects accessible to empirical observation and natural science.

① The phenomenal world includes everything as it is subject to our conditions of sense perception and ordered experience. These conditions include space, time, and strict causal connectedness according to necessary laws. From this standpoint, our actions are causally determined. If this were the only standpoint from which we could view ourselves, freedom would be impossible.

Kant, however, holds that nature is only a realm of appearances or things as they fall under the conditions of empirical knowledge. Everything in the phenomenal world, including ourselves, also has an existence in itself, not subject to those conditions. This other realm, that of things in themselves, cannot be known by us, but it can be thought, through our pure understanding, as a noumenal or intelligible world. Because we cannot know our noumenal selves, we cannot know whether they possess the freedom our phenomenal selves lack. Yet because space, time, and causal order are on Kant's theory necessary attributes only of the phenomenal world, there is no way to rule

★ out the possibility that the noumenal self is free. This leaves it open to us to postulate or presuppose noumenal freedom as the metaphysical condition of our moral agency, with which our consciousness of moral obligation acquaints us. Thus when we view ourselves as objects of empirical science, we regard our actions as causally determined and not free. On the other hand, when we assume the role of moral agents, we "transpose our-

Kant's Compatibilism

selves into members of intelligible world," regard our existence as not subject to the empirical conditions of space, time, or natural causality, and think of ourselves as free (G 453g 72e).

Kant's compatibilism is of a most unusual sort. The usual intention behind compatibilism is to deny that there really is any deep metaphysical problem about free will, to suppress the free will problem at the least possible metaphysical cost. Most compatibilists try to show how free choices may be comfortably accommodated in the chain of natural causes, so as to present a unified picture of ourselves as both moral agents and objects of scientific observation. Kant's compatibilism, however, is based on the aggressively metaphysical distinction between phenomena and noumena; far from unifying our view of ourselves, it says that freedom and determinism are compatible only because the self as free moral agent belongs to a different world from that of the self as natural object.) (Norman Kretzmann has commented to me that this may be likened to saying that a married couple is compatible, but only as long as they live in separate houses.)

this can't be literally correct

I venture to say that Kant's solution to the free will problem strikes nearly everyone who has ever studied it as thoroughly unsuccessful, a metaphysical monstrosity that gives us a far-fetched if not downright incoherent account of our moral agency. At any rate, that is what I used to think. I am still in partial agreement with those critics who find problems in reconciling Kant's account of our free agency with some of the more commonsense features of his own ethical theory. And I share with more orthodox compatibilists the hope that there may exist a simpler and less counterintuitive solution to the free will problem. Still, I am more impressed with the successes of Kant's theory in solving the problem it sets for itself than I am appalled by its shortcomings. As we shall see, the free will problem arises for Kant in a form in some respects idiosyncratic to Kant, and the problem for him is unusually acute. The causal determinism in which Kant believes is very strict, and the conditions for moral responsibility set by Kantian ethics are very strong. Further, Kant's moral psychology blocks the only simple and natural sort of compatibilist solution. What I find quite surprising is that

may be a reason to interpret him differently

Is the causal determinism so strict?

b/c sense of duty cannot be a natural cause

Kant does succeed in reconciling freedom and determinism in the context of his philosophy, and that Kant's compatibilism involves no greater revision of our commonsense view of our agency than it does.

KANT'S CONCEPT OF FREEDOM

dist. Kant distinguishes between 'transcendental' (or 'cosmological') freedom and 'practical' freedom. Transcendental freedom is a purely metaphysical concept, the concept of a certain sort of causality. Transcendental freedom is "the faculty of beginning a state spontaneously or from oneself [von selbst] (A533/B561). A transcendently free being is one that causes a state of the world without being subject in its causality to any further causes; that is, a "first cause," or an unconditioned beginner of a causal chain. Practical freedom, on the other hand, is free agency; it is what we ascribe to ourselves when we think of ourselves as morally responsible for what we do.

Negative characterization Regarding practical freedom, Kant distinguishes two concepts: a "negative" concept and a "positive" one. Practical freedom in the negative sense is "the independence of our will [Willkür] from necessitation through impulses of sensibility" (A534/B562). We are free in this sense if we are capable of resisting our sensuous desires, of acting contrary to their dictates or at least without having them as motives of our action. Kant apparently believes that nonhuman animals altogether lack the capacity to resist sensuous impulses, that the action of an animal is always the direct result of some sensuous impulse. When such an impulse occurs in an animal, it cannot fail to act as the impulse dictates. It has no faculty of countermanding the impulse or suspending action on it in order to deliberate and weigh different desires against each other. Kant calls the animal's will an *arbitrium brutum*, in contrast to the human *arbitrium sensitivum sed liberum*, which is affected by sensuous impulses but not necessitated by them (A534/B562).

This concept of practical freedom is "negative" because it de-

scribes the free will in terms of the way it does *not* operate. Kant's positive conception of practical freedom, on the other hand, describes the free will in terms of what it *can* do. Kant believes that unlike the brutes, human beings have the power to resist particular sensuous impulses and the power to deliberate about which ones to satisfy and how best to satisfy them. In addition, Kant ascribes to us the power to act from a wholly nonsensuous or a priori motive (to conform our will to a moral law self-given by our own reason). Freedom in the positive sense is thus the capacity to act from this nonsensuous motive, "the capacity to will a priori" (*VpR* 129g 104e). This positive sense of freedom is derived, of course, from Kant's moral theory. In general, we may be held responsible for our actions only to the extent that we have the capacity to do what morality commands of us. Kantian morality commands that we act according to a law we give ourselves a priori. Consequently, practical freedom consists in the capacity to be motivated by such a law; it is "the authentic legislation of pure and as such practical reason" (*KpV* 33g 33e).

The free will problem arises for Kant because he believes that practical freedom requires transcendental freedom and that there is no room in the causal mechanism of nature for a transcendentally free being (*A534/B562*). All causes in the natural world are sensible or empirical and all necessitate their effects. Consequently, if our actions were determined by them, then they would be necessitated by something sensuous (so that we would lack practical freedom in the negative sense) and there would be no room for a nonsensuous motive for our actions (so that we would lack practical freedom in the positive sense). Practical freedom requires that we be able to determine our actions entirely from within ourselves, through our own legislative reason. Natural causes, however, belong to an endless regressive chain in which there is no spontaneous or first cause. We can think of ourselves as practically free, therefore, only by thinking of our actions as subject to a transcendentally free cause lying outside nature.

Because practical freedom, the freedom required for moral responsibility, must involve the capacity to do what morality demands of us, it is obvious that any account of practical freedom

X
not the same
as acting on a
non-sensuous
motive

no, motivated
by respect for
such a law



Problems in
some of the
causes of our
actions are
non-sensuous
& sometimes
our actions are
not caused by
sensible intuition

— this hasn't
been shown
referenced

must depend on some account of the demands of morality. In Kant's philosophy morality has an especially close association with practical freedom. For Kant, morality is autonomy, it is the conformity of the will to a self-given law of reason. Acting morally, then, is not merely one thing among others that a practically free being can do, but it is the peculiar function of morality to actualize practical freedom.

It has seemed to some of Kant's critics, such as Henry Sidgwick, that an insoluble problem is involved in Kant's doctrine in that only autonomous or moral action actualizes practical freedom.² For in that case, apparently only moral or autonomous action is free action, while immoral or heteronomous actions are all unfree. If only actions motivated by pure reason are free, then actions motivated by sensuous desires must be necessitated by the mechanism of nature. Thus it seems that Kant is committed to saying that we are morally responsible only for moral or autonomous actions and that no one can be held responsible for immoral or heteronomous actions.

This problem is illusory, however. If Kant did hold that our immoral or heteronomous actions were determined by natural causes, then he would have to hold that these actions are unfree; that the will that performs them is an *arbitrium brutum*. But Kant's actual view is that since our will is free, our heteronomous actions are performed from sensuous motives without being necessitated by them. Even when we act from sensuous motives, we do so as having the capacity to be moved by the a priori law of reason; we act as practically free beings, and hence are responsible for what we do. If our actions were sensuously necessitated, as by natural determinism, then it would follow that they are all performed from sensuous motives. The converse of this does not hold, however. Not every action performed from sensuous motives must be naturally necessitated. The human will, according to Kant, is sensuously "affected" (*affiziert*) but not sensuously "necessitated" (*nötigt, necessitirt*). Sensuous motives, when we act on them, have for us a *vim impellentem* but not a *vim necessitantem* (A534/B562; VM 182).

²Henry Sidgwick, *Methods of Ethics* (New York, 1966), pp. 511–16.

It is true that when Kant describes practical freedom in the positive sense as "the legislation of pure practical reason," he seems to imply that only autonomous or moral actions are truly free actions. Careful attention to his language shows that for him practical freedom consists in the *capacity* for autonomous action and can exist even when this capacity is not exercised. Thus Kant describes practical freedom in the negative sense as "that property of [the will's] causality by which it *can* be effective independently of foreign causes" (G 446g 64e, italics added). Again, he says it is "not being *necessitated* to act by any sensuous determining ground," implying that a free being may be motivated by sensuous grounds without being necessitated by them (TL 226g 26e, italics added). Freedom in the positive sense, he says, is "the *power* [*Vermögen*] of pure reason to be of itself practical" (TL 213–14g 10e, italics added). And he defines the free will (*Willkür*) as "the will that *can* be determined by pure reason" (TL 213g 10e, italics altered). Further, "freedom of the will is the *capacity* to determine oneself to actions independently of a *causis subjectis* or sensuous impulses" (VpR 129g 104e, italics added). We are practically free and morally responsible for our actions whenever we act as having the power or capacity for autonomous action, whether or not that power is exercised or actualized in what we do. Kant's view is that except "in tenderest childhood, or insanity, or in great sadness which is only a species of insanity," we always do possess the power to act autonomously, even if we seldom exercise it (VM 182). Autonomous actions are "freer" than heteronomous ones in the sense that autonomous actions are an exercise of freedom—that is, they actualize the capacity in which practical freedom consists. In heteronomous actions, we fail to exercise our freedom (our power to act autonomously) insofar as we submit ourselves to the motivation of natural impulses. This is wholly different from action of an *arbitrium brutum*, which altogether lacks the power to determine itself independently of the sensuous impulses on which it acts.

Philosophers distinguish two senses of freedom that may be relevant to moral responsibility: the 'liberty of spontaneity' and the 'liberty of indifference'. Loosely speaking, we are free in the sense of spontaneity (if we are the cause of our own actions), and

free in the sense of indifference if we could have done otherwise than we in fact do.) Kant's theory of practical freedom appears to involve both spontaneity and indifference, at least in certain specific or qualified ways. Autonomous action is spontaneous in a very strong sense, because our acts have their determining grounds entirely a priori in our reason, and are wholly independent of all external causes. But even heteronomous action is spontaneous in quite a strong sense. Although heteronomous acts are motivated by sensuous desires, which Kant regards as 'foreign' to our rational will, they are not causally necessitated by such desires. As free agents, we act from sensuous motives only by incorporating them into our maxim, the principle adopted freely by our will. Practical freedom is always spontaneity because it requires transcendental freedom, the capacity to produce an effect without being determined by any prior or external cause.

Kant also ascribes a form of indifference to the human will. He even defines will (*Willkür*) generally as "a power to do or refrain at one's discretion" (TL 213g 9e). Clearly every one of our heteronomous actions could have been other than what it is, since it is performed by a being who has the capacity to act autonomously. Further, because our will is sensuously affected, the possibility always exists that we will fail to exercise our freedom. Hence every one of our autonomous actions also could have been other than what it is.

Kant, however, flatly refuses to define freedom in terms of this indifference. Freedom is the power to choose according to legislative reason. According to Kant, "freedom cannot be located in the fact that the rational subject can make a choice which conflicts with his legislating reason, even if experience proves often enough that this happens (TL 226g 26e). For Kant, freedom consists in a one-way difference, so to speak. That is, freedom consists in the ability to act autonomously even when we do not, but it does not consist in the possibility of acting heteronomously, even if this possibility always does exist for us. If an action is performed from sensuous motives, then it is a free action only if the agent could have acted instead on a priori motives. If a being acts on a priori motives, however, it may be free even if there is no possibility that it could have failed to act

Still, the
a priori
rational does not
necessitate the
action

as it does. A free will that altogether lacks this indifference is what Kant calls a "holy will," for example, the divine will:

[The capacity always to act according to reason must certainly be in God, since sensuous impulses are impossible to him.] One might raise the objection that God cannot decide otherwise than he does, and so he does not act freely but out of the necessity of his nature.... But in God it is not due to the necessity of his nature that he can decide only as he does. Rather it is true freedom in God that he decides only what is suitable to his highest understanding.
[VpR 132g 105–106e]

According to Kant's theory, practical freedom is the power to act from a priori principles. A being may have this power even if there is no possibility of failing to exercise it and hence no possibility of acting other than it does. Such a being is free, and therefore responsible for its inevitably rational acts.

Kant even goes so far as to say that "only freedom with regard to the inner legislation of reason is really a power; the possibility of deviating from legislative reason is a lack of power" (TL 227g 26e). Here again we must avoid the mistake of thinking that for Kant the failure to exercise our power to will autonomously is the same as the failure to have that power, hence the same as the failure to be free. What Kant means is that the possibility of deviating from legislative reason is not a power of any sort but is due instead to a certain sort of weakness or lack of power. This view is quite defensible. Not every possibility is a power. Some possibilities, in fact, are due to a lack of power. The human will has the power to act autonomously and this power is its freedom. If it acts heteronomously, it still has this power, but it has failed to exercise it. Such a failure is not due to a lack of freedom but rather is a failure of execution, a failure to exercise the freedom we have. And this failure is due in turn to a certain weakness or lack of power, leading to a lapse in exercising our freedom. Moral weakness, which makes it possible for our action to deviate from legislative reason, is a lack of power, but it is not a lack of that power in which freedom consists.

Consider a parallel case. An indifferent swimmer may have

X "weakness" was
but still an "ability"

Autonomy!
Action if
acting out of
respect for the
moral law or
simply acting
on the basis
of the moral
law!

the power to save himself if he falls into deep water. But he has no power, only a possibility, of drowning in the same eventuality. Although he has the power to swim, he may drown if he does not exercise that power effectively, due (say) to confusion or panic. In the latter case, his possibility of drowning is due to a kind of weakness or lack of power, though not to a lack of the power to swim. The indifference of the human will, on Kant's theory, is rather like the indifference of this swimmer. The swimmer has the power to swim, which he may exercise and save himself, or he may fail to exercise that power and drown, due to a weakness or lack of power leading to a failure of execution. The moral agent has freedom, the power to act autonomously, and may exercise that freedom by conforming one's will to the moral law, or one may fail to exercise it and follow sensuous desires, due to a weakness or lack of power in one's moral character leading to a failure of execution. The possibility of deviating from legislative reason is thus due to a lack of power and not to a lack of freedom.

We may now understand how the free will problem for Kant is different from the free will problem faced by most incompatibilists. Kant does *not* hold in general that freedom is incompatible with causal determination or even necessitation of the free being's actions. As we have just seen, Kant holds that a holy will is free even though its acts are necessitated, because they are necessitated from within by reason rather than by the sensuous impulses that are foreign to our rational nature. In this respect Kant's view is much closer to that of some compatibilists (such as Spinoza and Leibniz) than it is to that of standard, hard-line incompatibilists, who see freedom as incompatible with any sort of causation or necessitation.

Kant does not believe, then, that freedom is incompatible with causation generally, but only that it is incompatible with *natural causation*, with the sort of causation found to act on the will within the realm of appearance. He believes this for two overlapping reasons, one of them derived from his moral philosophy and the other from his moral psychology. First, Kant believes that the moral motive, action on which is alone an exercise of freedom, is an *a priori* motive, and therefore one that cannot

There may be
other causes
than natural
causes in the
realm of appearance

be given to us through nature or the causes acting in nature. Second, he holds the complementary belief that the only natural causes that do act on the will are sensuous impulses, motivation by which precludes motivation by reason. In short, the free will problem arises for Kant because he is a thoroughgoing psychological ~~hedonist~~ about all the natural causes that might act on our will. He holds that empirical psychology is excluded in principle from understanding all rational deliberation and all action on the motive of reason. This means that the free will problem would not arise in the same form for someone who is less of a crude Benthamite about the empirical psychology of motivating and allows for the possibility of accounting for rational deliberation and action through natural causes. A standard compatibilist response to the free will problem in its Kantian form seems open to any such person.

★

He hasn't said
that these causes
are pleasure, just
that they are
sensuous

FREEDOM AND NATURE

Let us now turn to Kant's attempt to solve his problem, to reconcile freedom with the mechanism of nature given the strictures of his moral philosophy and moral psychology. First we must clarify what Kant does and does not intend to establish. Kant does not pretend to prove that we are free or to provide arguments in favor of any speculative doctrine he believes must be true if freedom and determinism are to be compatible. No such doctrines, he holds, could be either provable or disprovable. What Kant means to show is only that, as far as we can prove, freedom and determinism may be compatible. Kant is fond of forensic analogies. Let us use one here. Kant's role regarding freedom is somewhat like a defense attorney's role regarding his client. Because practical freedom is presupposed by morality, we may assume that freedom is innocent until proven guilty, that the burden of proof lies on those who would undermine our moral consciousness by claiming that we are not free. In confronting the prosecution's evidence against his client, a defense attorney may exploit this fact about the burden of proof

by concocting a plausible theory explaining the allegedly incriminating evidence. The attorney need not show that his theory is the correct one, only that it has enough plausibility to introduce a reasonable doubt concerning the prosecution's theory. Likewise, in defending our freedom, Kant concocts a metaphysical theory which, if true, saves our practical freedom despite the fact that our actions are determined by natural causes. Kant does not need to show that this metaphysical theory is correct; indeed, he frankly admits that he cannot. But neither, he claims, can the opponent of freedom refute the theory. Kant rests his defense of freedom on his claim.

The problem Kant faces is formidable. As we have just seen, Kant holds that the human will has freedom, which is the power to resist sensuous impulses and to act from nonsensuous motives.

At the same time, Kant also holds that human actions fall under the unexceptionable law of natural causality:

All the actions of man in appearance are so determined . . . according to the order of nature that if we could investigate all the appearances of his will as to their grounds, then there would not be a single human action we could not predict with certainty and be able to know as necessary from its preceding conditions. [A550/B578]

Kant hopes to save himself from an open contradiction here by distinguishing two standpoints for viewing our actions, each involving a different sort of cause to which the action may be attributed. It is possible, he says, "to regard an event on the one side as a merely natural effect, yet on the other side as an effect of freedom" (A543/B571). As an appearance or phenomenon, the human self plays a natural causal role in the production of its actions, which are events in nature and appearances as well.

Every efficient cause, however, must have a *character*, i.e. a law of its causality, without which it would not be a cause. And then we as subjects in the world of sense have first an *empirical* character, through which our actions as appearances stand in thoroughgoing

Kant's Compatibilism

connection with other appearances according to constant laws of nature. [A539/B567]

We also have an existence in itself as noumenon or intelligible subject, outside the conditions of sensible nature, unknowable but still thinkable through pure understanding. In this subject we "must make room for an *intelligible character* through which it is the cause of those same actions as appearances, but which is not itself an appearance and stands under no conditions of sensibility" (A539/B567). As effects of our empirical character, our actions are necessitated by natural causes and hence unfree. But as effects of our intelligible character, it is possible that the same actions are produced by a transcendently free cause, which is not necessitated by anything sensuous and is capable of autonomous or a priori volition.

No. Empirical
Character is
simply a causal
law or set of laws

It is not immediately clear how this theory is supposed to reconcile freedom and determinism, or even how it can be self-consistent. If our actions are indeed causally determined by natural events, then they are apparently necessitated by sensuous impulses acting on our empirical character. From this alone it seems to follow that our actions are unfree. How can anything that might be true about our actions from another standpoint render these same actions free? How can it remove the necessitation of our actions by sensuous causes or restore to us the capacity of acting from a priori motives, which this necessitation appears once and for all to preclude?

Yes.
If only sensuous
impulses can be
natural causes

Two problems arise here; one concerns the liberty of spontaneity and the other the liberty of indifference. First, practical freedom is the capacity to will autonomously, to act from a priori motives. How is it possible for us to have this capacity if a natural or empirical cause can always be cited for what we do? Second, moral responsibility requires that for any heteronomous action we perform, we must have the capacity to act other than we in fact do. How is this possible if every one of our actions is in principle predictable and knowable as "necessary from its preceding conditions"?

The cause would
not necessitate the
effect if our
empirical char.
had been diff.

K didn't say
empirical

I believe Kant is aware of both problems and attempts to solve them. Let us consider each attempt in turn.

EMPIRICAL CAUSALITY AND INTELLIGIBLE CAUSALITY

Regarding the problem of spontaneity, I believe that a careful consideration of Kant's texts reveals that according to the theory he is proposing, our free actions are *never* produced by natural causality in the way they would have to be if only natural causality pertained to them. Kant holds that the two sorts of causality do not merely exist side by side but that, at least in the case of human actions, phenomenal causality is grounded in noumenal causality. He asks:

Is it not possible that although for every effect in appearance a connection with its cause according to natural laws is indeed required, this empirical causality itself, without in the least interrupting its connection with natural causes, can be an effect of a causality which is not empirical but intelligible? [A544/B572]

More specifically, Kant holds that although our actions are causally determined in time by our empirical character and other natural events, this empirical character itself is the effect—or, as he also says, the "sensible schema" (A553/B581)—of the intelligible character, which is freely determined by us outside empirical conditions. Kant says that "reason has causality in regard to appearance; its action can be called free, since it is exactly determined and necessary in its empirical causality (the mode of sense). For the latter is once again determined in the intelligible character (the mode of thought)" (A551/B579).

Empirical causality regarding human actions is an "effect" of intelligible causality, which (on the theory Kant is proposing) is transcendently free. Hence empirical causality, on this theory, does not involve the sensuous necessitation of actions, as it appears to do when we ignore its intelligible ground. Practical freedom, says Kant,

Kant's Compatibilism

presupposes that . . . an action's cause in appearance is not so determining as to preclude a causality lying in our will, a causality which, independently of these natural causes and even contrary to their force and influence, can bring about something determined in the temporal order according to empirical laws, and thus can begin a series of events wholly of itself. [A534/B562]

→ a particular causal influence!

Again, "If appearances were things in themselves, then freedom could not be saved. For then nature would be the complete and self-sufficient determining cause of every event" (A537/B565). Kant's theory apparently holds that because appearances are not things in themselves (nature is *not* the complete and self-sufficient cause of events, at least not of human actions) Rather, the complete and self-sufficient cause of actions is our free will, located in the intelligible world. Nature, in the form of sensuous impulses, enters into the production of our actions only insofar as we freely permit sensuous motives to be substituted for a priori rational principles in determining our choices. The free causality of the noumenal self, Kant says, "is not merely a concurrence, but complete in itself, even when sensuous incentives are not for it but wholly against it" (A555/B583).

+ this world
+ still unconditioned

It is tempting to describe Kant's theory by saying that for this theory the natural or empirical causes of actions, their causes in the world of appearance, are not real causes but only apparent causes; furthermore, that on this theory everything in the phenomenal world goes on, by a sort of preestablished harmony, just as if our actions were caused by antecedent natural events, but in reality their causes lie outside nature altogether, in a free will hovering above nature in the intelligible world. I believe, however, that Kant would reject this description as a distortion or caricature of his theory. Kant's principle of empirical causality says that every event in time is determined by antecedent events according to necessary laws. For Kant every human action does conform to this principle. This conformity to natural laws entitles the empirical events upon which our actions necessarily follow to be called the empirical causes of those actions: nothing more is, or could possibly be, required for them to deserve that title. They are not, therefore, merely apparent causes, but the real

metaphorical language
Against necessary behavior

causes of our actions insofar as they fall under the mechanism of nature. The antecedent events are not "complete" and "self-sufficient" causes, however, because the causality of human actions can be viewed from another standpoint, that is, as the effects of freedom.

Although Kant's position here is thoroughly self-consistent, it reveals something noteworthy and possibly suspicious about his conception of causality. The concept of cause is bound up with that of causal efficacy, a notion David Hume variously, and aptly, called a cause's "power," "energy," "force," or "that very circumstance in the cause, by which it is enabled to produce the effect."³ Regarding this notion philosophers have exhibited two general tendencies. One tendency, which we may call the Aristotelian, is to treat causal efficacy as a property of substances or agents and to regard it as a primitive notion, unanalyzable but built into our basic understanding of how things work. The other, which we may call the Humean tendency, is to treat causal efficacy as a property of events or states of the world and to try to analyze it in terms of some relation between these events allegedly simpler or more accessible to our knowledge than that of causal efficacy itself.) The two leading candidates for such an analysis have been the relation of constant conjunction between similar events and the conformity of events to natural laws governing their temporal order.

Kant does not opt unambiguously for either tendency. He wants to treat causality in the sensible world as the conformity of events to laws determining their order in time. Kant's disagreement with Hume about whether these laws are knowably necessary is a disagreement between two philosophers who both follow the Humean tendency regarding natural causality. Kant, however, wants to reserve for himself a more Aristotelian notion of causal efficacy, to be ascribed to free agents as members of the intelligible world. If we agree that Kant's theory at this point is coherent, then it seems to provide a counterexample to the extreme Humean view that we can conceive of causal efficacy

³David Hume, *Enquiry Concerning Human Understanding* (Oxford, 1979) section 7.

only as some feature of the temporal order of events. For the theory says that our actions are determined in the temporal order by natural laws, but the causal efficacy responsible for them does not lie in the temporal series at all, but outside it in the intelligible world. If we can imagine this at all, then it suffices to show both that we can conceive of events as lawfully ordered in time without attributing causal efficacy to those events upon which others lawfully follow, and that we have a notion of causal efficacy which has nothing whatever to do with the temporal order of events and the laws governing it. For this reason, stubborn Humeans may reject Kant's theory as unintelligible and his solution to the free will problem along with it. I will not pursue this issue here. But we have seen that if we assume Kant's views about causality are coherent, we must agree that his theory succeeds in showing how natural causality could be compatible with the spontaneity of the human will in its intelligible character.⁴

TIMELESS AGENCY

There remains the problem of indifference, of showing how we can have the capacity to act autonomously even when we do not so act, and even when our heteronomous actions are causally determined to take place. Kant admits that insofar as we conceive ourselves as empirical agents, who cause our actions through the

⁴I am inclined to say that Kant's refusal to opt exclusively for the Humean concept of cause sheds light on another of the vexed questions of Kantian metaphysics: the causal relation between noumena and phenomena generally. Kant holds that things in themselves cause appearances, or that they appear to us by exercising a causal influence on our sensibility. For two centuries critics have charged Kant with inconsistency or incoherence for holding this doctrine. The critics charge that on Kant's own showing causal relations can obtain only between temporal events, so that it is inconsistent for Kant to regard things in themselves (which are timeless) as causes. It is certainly true that for Kant all empirically knowable causes hold between events in time, since such causes are all Humean. Kant means to allow us through pure understanding to *think* of Aristotelian causes as the noumenal ground of phenomenal objects. The pure category of cause and effect (or ground and consequence) seems to be for him neither exclusively Aristotelian nor exclusively Humean. Here again, Kant is thoroughly self-consistent, although it is easy to see why his critics say he is not and hard to blame them for perceiving something fishy about Kant's doctrine at this point.

temporal flow of events, we cannot conceive of ourselves as able to do other than we do. Every one of our actions is strictly determined by the temporal series that precedes it, and this series reaches far into the past, before our birth. "Since the past is no longer in my power, every action which I perform is necessary because of determining grounds which are not in my power. This means that *at the point in time when I act I am never free*" (*KpV* 94g 98e, italics added). Time, however, is for Kant only a form of sensibility; only as phenomena or appearances are we necessarily in time. As noumena or things in themselves we are subject neither to time nor to the law of causality which goes along with it. As free agents, according to the theory Kant proposes, we are timeless beings.

Kant here presupposes another controversial metaphysical doctrine, that of timeless eternity. The attribute of timelessness, which some theologians ascribe to God alone, Kant's theory must ascribe to every one of us insofar as we are transcendently free. Some philosophers doubt that we can make sense of the notion of timeless eternity. Some object especially to the idea that a timeless being might be the cause of temporal events, and in particular they object to the idea that it can be thought of as causing events that occur at different times. Clearly, if we cannot form a coherent notion of timeless eternity or a coherent account of a timeless being causing events occurring at different times, then Kant's solution to the problem of freedom is untenable. Once again, however, I will not try to settle this difficult metaphysical issue here. The point I want to investigate is whether, granted that such an account can be coherently formulated, Kant is capable of reconciling natural determinism with the indifference necessary for moral agency.

How is the timelessness of the noumenal self supposed to safeguard our ability to do otherwise than we in fact do? In order to answer this question, we must look more closely at Kant's metaphysical theory of the case, and in particular at the way in which our timeless existence as noumena is supposed to produce our actions, which unfold in the course of time. In the moral subject's noumenal existence, Kant says,

Nothing precedes the determination of his will; every action and, in general, every changing determination of his existence, according to the inner sense, even the entire history of his existence as a sensuous being, is seen in the consciousness of his intelligible existence as only a consequence, not as a determining ground of his causality as a noumenon. From this point of view, a rational being can rightly say of any unlawful action which he has done that he could have left it undone, even if as an appearance it was sufficiently determined in the past and thus far was inescapably necessary. For this action and everything in the past which determined it belong to a single phenomenon of his character, which he himself creates.
[KpV 97–98g 101e]

Kant's theory seems to be the following. Events in time follow a necessary order, as determined by their natural causes. A particular timeless choice of my intelligible character affects the natural world by selecting a certain subset of possible worlds, namely, those including a certain moral history for my empirical character, and determining that the actual world will be drawn from that subset of possibilities. For each such choice there is an almost endless variety of ways in which I might have chosen differently, and endless variety of possible empirical selves and personal moral histories I might have actualized. Of every one of my misdeeds it is true that I would have left it undone had I made a different timeless choice. Hence it is in my power to leave any misdeed undone, despite the fact that in the actual world it follows inescapably from what preceded it in time.

When Kant says of my unlawful action that it and its causal determinants belong to my character, he may intend to claim that my timeless choice affects the course of the phenomenal world merely by affecting the constitution of my empirical character, which is of course an important causal factor in what I do. This would bring Kant's theory into line with the common-sense idea that we are responsible for our actions because we are responsible for our characters. An obstacle to this way of squaring Kant's theory with common sense is that on his theory the phenomenal effects of my timeless choice appear to extend far beyond the constitution of my empirical character. Indeed,

yes - by
determining
the temporal
natural laws that
apply to my

*Where does
Wood get this
?*

my empirical character is itself causally determined by preceding events, and ultimately by events very remote in time from my life and actions. If I am responsible for my character, then my timeless choice must affect the whole course of the world's history, insofar as history includes the actions I in fact perform. Ralph Walker even maintains Kant's theory has the monstrous consequence that I am morally responsible for everything that happens in the course of the actual world, since this world as a whole results from my timeless noumenal choice: "I can be blamed for the First World War, and for the Lisbon earthquake that so appalled Voltaire. Gandhi is no less guilty than Amin of the atrocities of the Ugandan dictator."⁵

In rescuing Kant from these supposed consequences of his theory, it may help to keep in mind that on Kant's theory my intelligible choice is supposed to impinge on the course of the world not by directly selecting a certain history of actions for me. Rather, what my intelligible choice fundamentally decides is my empirical character, the kind of person I will be, or as Kant puts it in the *Religion*, the "fundamental maxim" on which I will act (*Rel* 31g 26e). Even if my choice somehow issues in a world containing the First World War, the Lisbon earthquake, and the deeds of Idi Amin, it seems reasonable to hold me morally responsible only for those events which must belong to the actual course of things because I have the empirical character or fundamental maxim that I do. Kant must admit that on his theory this may include events that happen at places and even times remote from my life history in the temporal world. Yet Kant can reply that because in principle we know nothing about how our timeless choices operate on the temporal world, it must be impossible for us to say with confidence which events these may be. It seems open to Kant to suppose that they correspond to those events for which we normally regard ourselves as morally responsible. We must keep in mind that the purpose of Kant's theory is not positively to establish that we are free or morally responsible for our actions and their consequences, but only to suggest one possible way in which the moral responsibility

⁵Ralph Walker, *Kant* (London, 1979), p. 149.

we ascribe to ourselves may be reconciled with the mechanism of nature. We altogether misconstrue the status and function of this theory for Kant if we try to use it to justify or preempt moral common sense concerning such matters as the empirical scope of our moral responsibility.

FATALISM

But here we must face another objection to Kant's theory not too different from the preceding one. Although Kant does permit us to say of our unlawful actions that we could have left them undone, it may appear that his theory forces on us a certain fatalism about our character and actions at odds both with common sense and with the spirit of Kant's own moral philosophy. As we have already seen, Kant allows that if we had enough knowledge of someone's character and circumstances, "his future conduct could be calculated with as much certainty as a solar or lunar eclipse" (*KpV* 99g 103e). Suppose, then, that I have enough knowledge about myself to calculate that I will perform a certain unlawful action, for instance, that I will tell a certain malicious lie about an acquaintance against whom I bear an envious grudge. On Kant's theory it is true that I can avoid telling this lie, in the sense that I could have timelessly chosen a different intelligible character from my actual one, resulting in a different empirical character and a different moral history for myself. Yet it appears I already know enough about myself to be certain of my timeless choice regarding this particular action. For I know I will tell the lie and I know this lie is inescapably determined by past events which have already occurred. Hence I know to that extent which possible world I have timelessly chosen to actualize. My future action of telling the lie must therefore seem to me like a past action, or rather like the inevitable future result of a past action, which I can see coming but can no longer prevent. My malicious lie is an action I could have avoided (by choosing a world with a past different from the actual one) but it seems to be an action from which I am

this can't be right

now powerless to refrain. As I now view myself, it seems that both my character (as a malicious liar) and my future act of lying are something fated for me. Perhaps I am in some sense to blame for them, but I cannot view myself as now able to alter or to avoid them.

This fatalism is obviously incompatible with the spirit of Kant's moral theory. But one measure of its appeal as an interpretation of Kant's theory of freedom is the fact that it was adopted, not only as an interpretation of Kant, but even as the profound truth of the matter, by Arthur Schopenhauer. Citing Kant's theory of the intelligible and empirical characters, Schopenhauer alleges that it supports his own thesis that everyone's conduct is inescapably determined for good or ill by an innate and unalterable moral nature. According to Schopenhauer, the merit of Kant's theory is that it shows that I am wholly responsible for what I do, despite the fact that it is false to say that I could have done otherwise, except in the sense that in the same circumstances, a different person, with a character different from mine, would have acted differently.⁶

I believe that Schopenhauer's interpretation of Kant is based on some clearly fallacious reasoning from Kant's theory. Insofar as a person's intelligible character is timeless, it apparently must also be immutable. Certainly, however, Kant intends his theory to be compatible with every conceivable state of affairs concerning the constancy or alterability of our empirical character through the course of a lifetime. Presumably this theory is that for every imaginable course of conduct in the phenomenal world, there is a timeless choice of an intelligible character that would yield that course of conduct. Hence there are some such choices whose results in the world of appearance involve changes in empirical character, drastic conversions from evil to good, or sudden degenerations from good to evil. For Kant, whether there are such changes in fact cannot be settled a priori on metaphysical grounds.

Kant also has a reply to the charge that his theory implies we

⁶Arthur Schopenhauer, *The World as Will and Representation*, tr. E. J. F. Payne (New York, 1958), 1:286–307; *On the Freedom of the Will*, tr. K. Kolenda (New York, 1960), pp. 26–64, 81–83; *On the Basis of Morality*, tr. E. J. F. Payne (Indianapolis, 1965), pp. 109–115.

no longer have the power to alter our future conduct. Kant does hold that by knowing enough about our character and situation we can predict with certainty what we will do. Yet it does not follow from this alone that we lack the power to do otherwise. Consider once again my malicious lie. I know how strong my grudge against the victim is, and I know myself well enough to be quite certain that I will persist in my resolve to slander him, that no sense of shame, no pity for my victim, still less any moral scruples, will dissuade me from carrying out my sinister intention. I can predict with absolute certainty that I will tell the malicious lie. But however certain I may be, I am not in the least inclined to think that it is not in my power to refrain from lying when the time comes. On the contrary, although I am certain that I will tell the lie, I am no less convinced that I still can refrain from telling it. It is misleading to express my conviction, as shallower compatibilists sometimes do, by saying that I could refrain from lying *if* I wanted to refrain. For that might suggest that, given my actual wants, I do *not* have the power to refrain. This, however, is exactly what I do *not* believe. Rather, what I believe is that although I know beyond a shadow of a doubt that I *will* tell the lie because I know I want to tell it, nevertheless I still *can* refrain from telling the lie.

Why do I have this belief, despite my certainty about how and why I will act? The belief, I suggest, rests on two other beliefs I have about myself and my future action. First, I believe it depends on me whether I will tell the lie or not. Second, I believe my influence on the situation will become effective regarding the lie only at the time the lie is actually told and not before. These are the reasons I believe that it is now still in my power whether I will lie or not. If Kant's theory can accommodate these two beliefs about myself, it can also accommodate my belief that it is still in my power to refrain from telling the malicious lie I know I am going to tell.

Kant's theory *can* accommodate both beliefs. As we have seen, this theory does show how my actions all depend on me alone, despite the fact that they follow causal laws and are predictable according to them. Kant also holds that my free choice itself becomes effective regarding a given action only at the time the

action is performed. "Every action," says Kant, "irrespective of the time relation in which it stands to other appearances, is the immediate effect of the intelligible character" (A553/B581). It is an illusion to suppose that my timeless choice of an intelligible character is located so to speak at the beginning of time, so that it operates *through* the series of natural causes. Rather, we must treat our timeless choice as spontaneously determining each individual act as that act occurs in time, so that in judging it "we presuppose that we can wholly set aside how [the agent's previous course of life] may have been constituted, and regard the past series of conditions as not having happened, but regard this act as wholly unconditioned with respect to the previous state, as if with this act the agent instigated a series of consequences wholly of itself" (A555/B583).

Kant's theory at this point resembles Boethius's resolution of the problem of human freedom and divine foreknowledge. Boethius argues that because God is timeless, God's knowledge of what we will do in the future is not literally foreknowledge. Focusing on the certainty and immediacy of this knowledge, we do best to say that God knows perfectly what we do *simultaneously* with our future action, much as I might know an event that is transpiring before my very eyes. Just as my present knowledge does not predetermine or constrain what is happening, so God's perfect and immediate knowledge of what we will do does not compromise our freedom.⁷ In a similar manner, Kant's theory says that our timeless choice does not predetermine our actions but has its influence immediately on each of them and should be considered simultaneous with each act as it occurs in the temporal order. Once again, some philosophers maintain that it is incoherent to claim that a single timeless choice can be simultaneous with each of a number of different events occurring at different times. This problem (which I do not believe is insoluble) is built into the very notion of a timeless being exerting influence

⁷Boethius, *Consolatio philosophiae*, *Corpus Christianorum* (Turnholti, 1957), Lib. 5, Prose 4–Prose 6.

on temporal events.⁸ Kant's idea that our timeless choice is simultaneous with each of our acts raises no *new* problems not already involved in the notion of timeless agency, and it follows closely one main line of traditional thinking about how timeless agency might operate. Kant's theory, therefore, does not commit him to fatalism about our future actions.

Of course, it is still obvious that this theory does not leave intact our commonsense conception of our free agency. As countless critics of Kant have observed, we surely do think of our moral agency as situated in time. We suppose that our free choices are made in the temporal flow, reacting to the course of events as it unfolds. We believe that we are free "at the point in time when we act," and not timelessly, as Kant's theory requires.

Timeless agency also forces other revisions in our self-conception as moral agents not easily harmonized with Kantian ethics. For example, it makes nonsense of the goal of moral improvement or moral progress, literally understood, a goal to which Kant often attaches considerable importance. On Kant's theory, only our external actions occur in time; our freedom, to which alone true moral worthiness or unworthiness pertains, is timeless, and hence incapable of literally changing for better or worse.

Further, although Kant's theory may allow us to conceive of ourselves as acting on the world at different times, it does not seem that it can consistently allow us to conceive of ourselves as acting *in* time, or producing events *within* and *through* intervals of time. Thus Kant's theory cannot permit us to conceive of ourselves as *trying* or *striving* to produce a certain result over a period of time. Kant's theory need not deprive us of every conception of ourselves as trying or striving, since there may be a notion of trying or striving that is instantaneous and simultaneous with the result striven for. (The old notion of a mental 'volition' simultaneous with every voluntary action might be one

⁸The problem is well stated by Anthony Kenny, "Divine Foreknowledge and Human Freedom" in Kenny, ed., *Aquinas* (Garden City, N.Y., 1969). It is equally well resolved, in my opinion, by Norman Kretzmann and Eleonore Stump, "Eternity," *Journal of Philosophy* 78 (1981), 429–58.

such conception of ‘trying’.) But Kant’s theory does deprive us of the idea that we may produce a future result by continuously striving to produce it throughout the interval of time which now separates us from it. For instance, there is no place in Kant’s theory for the idea that I may resist some passion or inclination of mine tomorrow by struggling with it today, and by striving throughout the day to purify my motives and fortify myself for the crucial hour of decision. The problem is not that I cannot imagine myself having all the thoughts and performing all the actions that I think of as part of this process, for certainly I can. The problem is that I cannot think of them as connected parts of an exercise of agency through time. I can only think of them as results or products of (timeless) agency, and not as the actual exercise of it. In time, there are only *facta*; yet trying or striving is not a *factum* but a *facere*. It is this exercise of agency Kant’s theory will not allow me to conceive as a temporal process.

The absence of trying or striving is of course no problem for a timeless God, whose omnipotence and moral perfection presumably obviate the need for trying or striving of any sort. The absence surely is a problem for dedicated Kantian moral agents, however, who must think of themselves as struggling constantly with their unruly inclinations and striving throughout their lives to make the idea of duty the sufficient motive of every action. In some writings after 1793, Kant seems to recognize that his theory cannot accommodate moral striving or moral progress, literally speaking. And he seems to want to employ the notion of our noumenal “disposition” or “attitude” (*Gesinnung*) as a sort of timeless analogue or substitute both for moral striving and moral progress. In these passages, Kant seems to be saying that we think of ourselves as striving and morally progressing in time only because we as temporal beings can form no positive notion of timeless eternity: “For then nothing remains for us but to think of an endlessly progressing change in the constant progress toward our final end, through which our *attitude* remains always the same. (But our attitude is not, like this progress, a phenomenon; rather it is something supersensible, and so does not alter in time)” (*EaD* 334g 98e). In the moral life, then, as in Goethe’s heaven, *alles Vergängliche ist nur ein Gleichnis*—“Everything tran-

sitory is only a parable." Temporal striving and moral progress are the moving images of our eternal moral attitude, which we cannot conceive directly but to which we can relate only through such temporal images or parables. Kant thus hopes to preserve our ordinary experience of moral struggles in time and moral progress through time not as literal truth, but only as the best way we have of representing to ourselves a truth we cannot directly experience or literally comprehend. Nevertheless, Kant must not pretend to deny that his theory requires staggering revisions in our commonsense conception of our agency, even in those features of his theory very dear to Kant himself.

In assessing Kant's compatibilism, it may help to remind ourselves that his theory of timeless agency is put forward only as a means of exploiting the burden of proof in the free will problem, which falls to those who would show that freedom is incompatible with determinism. Kant is not positively committed to his theory of the case as an account of the way our free agency actually works. Indeed, Kant maintains that no such positive account can ever be obtained. Kant does not pretend to know how our free agency is possible, but claims only to show that the impossibility of freedom is forever indemonstrable. If what bothers us about Kant's theory is that it seems too farfetched and metaphysical, then it may help at least a little to realize that once the theory has served Kant as a device for showing that freedom and determinism cannot be proven incompatible, he is just as content to dissociate himself from it and adopt a largely agnostic position on the question how our freedom is possible.⁹

Yes, but it
would be likely
for him to
speculate only in
ways he thought
were plausible for
ways he agreed with

⁹It may be argued that the implausibility of Kant's theory makes it unsuitable to his purpose, in the following way. According to the analogy drawn earlier, Kant may be likened to a defense attorney offering a theory of the evidence with a view to creating a reasonable doubt concerning the prosecution's theory. In a courtroom not just any logically possible theory of the evidence will do: the defense attorney's theory must be plausible enough to create a reasonable doubt. Some may believe Kant's theory of noumenal causality and timeless agency is too farfetched to pass this test. Kant's reply to this sort of objection is quite clear, however. In matters of transcendent metaphysics, questions of conjecture or probable opinion never arise: "In this species of investigation, it is in no way allowed to hold *opinions*. Everything which looks like a hypothesis is a forbidden commodity; it should not be put up for sale even for the lowest price, but should be confiscated as soon as it is discovered" (Axiv). This case is unlike the typical

On the other hand, we are justified in expressing some discontent over the fact that the best theory Kant can come up with is one that involves such radical revisions in our conception of free agency. At any rate, this shortcoming makes it difficult to credit the advertisement frequently given for Kantian morality, that it is a moral philosophy faithfully representing the moral life as the ordinary agent experiences and lives it. As with many advertisements, this one calls our attention only to the more palatable and wholesome ingredients in the product, and carefully avoids listing the artificial ingredients with ugly names which, though necessary to keep the product from spoiling, may render it much less appealing.

Before we dismiss Kant's solution, we must take a long, sober look at his problem. The fact is that the free will problem is an old and intractable one. It is bound to be especially so for a strict determinist whose moral philosophy and moral psychology require as a condition of responsibility that I be capable of actions whose motivation lies altogether outside my natural, sensuous being.

Even those who do not face the free will problem in its Kantian form must confront the difficulties raised by more standard incompatibilists. We may doubt that any solution to these difficulties exists that does not force us to abandon common sense in some way, either to alter the moral judgments we make or to revise our commonsense conception of our agency. In the end, solving the free will problem may not be a matter of "saving common sense" (for that may be quite hopeless). Rather the solution may be a matter of saving as much of it as we can, and especially of saving those parts of it which matter most to us. I believe Kant saw the situation in this way, and I suggest we may assume that he decided that the temporality of our agency is the

court case in that there is no question here of deciding whose theory is the more probable. In matters of metaphysics, any theory which can be neither proven nor disproven with apodictic certainty must count as equally probable. If Kant's theory of the case cannot be strictly disproven, then the contentions of the opponent of freedom cannot be proven. Because the burden of proof is on the opponent of freedom in this case, it follows that the only verdict we can render is to acquit freedom of the charges brought against it.

necessary ransom that must be paid to the free will problem if our high vocation as moral agents is to be preserved.

Others, of course, may have different worries and priorities and may wish to negotiate a different settlement. Nonetheless, they should not suppose they can get something for nothing. Moreover, they should evaluate Kant's solution in terms of his own problem. From this standpoint, we may conclude that unless we are prepared to argue for the positive incoherence of certain doctrines which constitute Kant's theory of the case—doctrines about phenomena and noumena, about Humean causality in nature and Aristotelian causality outside it, or about timeless agency and its temporal effects—then we must judge Kant's compatibilism a success in solving the problem it sets itself.