

AN ACTION-PLAN INTERPRETATION
OF PURPOSIVE EXPLANATIONS OF ACTIONS

This paper deals with the interpretation of Purposive ("teleological", "reason") Explanations of Actions (PEA's). My first task is the circumscription of the topic.

I

I shall take PEA's to be canonically formulated by the 'in order to' locution.

- (1) He got up early in order to get some yard work done.
- (2) He opened the refrigerator in order to get some beer.

There are, of course, other ways of providing essentially the same information as (1).¹

- (3) He got up early because he wanted to get some yard work done.
- (4) He got up early because he thought it would be a good time to get some yard work done.
- (5) He got up early because he likes working in the yard in the early morning.
- (6) He got up early because he had to get some yard work done.

I shall restrict this discussion to the purposive explanation of *intentional* actions. I am not, however, restricting intentional actions to those done in accordance with a prior intention. I include all those actions that involve a carrying out of present intentions, intentions one has in acting. I shall also restrict the discussion to overt actions, actions that essentially involve peripheral bodily movements. I do not deny that there are other intentional actions, e.g., rehearsing a speech in my mind; but I will not treat them here.

The dominant interpretation of PEA's on the current English language philosophical scene is the Causal Explanation Interpretation (CEI). The CEI takes its lead from the fact that PEA's can be formulated in the way exemplified by (3) and (4). Whereas (1) and (2) might suggest that a PEA appeals to "final causality", takes a future goal to explain the action that is "directed" to that goal, (3) and (4) suggest rather that the PEA does its explanatory work by citing antecedent psychological states that could plausibly be regarded as among the causes of the action. More specifically, the idea is that in giving a PEA we are, explicitly or implicitly, citing a want (using this term in the widest possible sense as ranging over all "pro-attitudes") for a certain state of affairs (getting yard work done, having beer, or whatever), and a belief that performing the action in question will (at least probably) contribute, in one way or another, to the realization of that state of affairs, and claiming that this want and belief are causes of the action.² ('Cause' here is not to be understood as 'sufficient cause' but rather as 'causal contributor'.) Normally we don't explicitly mention both the want and belief, and sometimes neither will be mentioned as such, as in the 'in order to' form, but, with help from the context, the relevant want and belief can be reconstructed from what is said. Here are some examples.

- (7) He turned the key to start the motor.
 Want – for the motor to be started
 Belief – that turning the key would result in the motor's being started.
- (8) He left the meeting because it was 8:00.
 Want – to be somewhere other than the meeting at or shortly after 8:00.
 Belief – that to be there then it was necessary to leave the meeting by 8:00.
- (9) He said 'That's very nice' because he realized that he should be polite.
 Want – to be polite
 Belief – that saying 'That's very nice', in those circumstances, was being polite.

Note that there are various ways in which an action can contribute (and can be believed to contribute) to the realisation of a state of affairs. It

can bring about the state of affairs (7); it can put the agent in a position to bring about the state of affairs (1) and (2); it can, given appropriate circumstances, constitute the state of affairs (9); and so on.

The CEI needs a great deal of further refinement, but that is outside the scope of this paper, since our concern here is rather to suggest an alternative. However there is one point that must be specified since it is crucial for one of our main objections to the theory. The causes of an action that the CEI takes a PEA to cite are *antecedents* causes, causes that obtain before the action occurs and that are followed by the action. Some of those who have read or heard earlier versions of this paper have expressed scepticism about this; and it must be confessed that causal theorists are less explicit on the point than one may wish. They do not, typically, address the issue straightforwardly. Nevertheless one can find a variety of indications that this is the way they are thinking of the matter. To illustrate this I choose one of the more extended presentations of the CEI, Alvin Goldman's book, *A theory of Human Action* (1970). Goldman's most extended formulation of the CEI is not wholly explicit on the point.

Thus the statement that *S* flipped the switch *in order to* turn on the light implies more than that *S had* the indicated want and *had* the indicated belief. It also implies that his having this want and his having this belief *caused*, or *resulted in*, his flipping the switch. Such an explanation not only implies that he *had* an action-plan that included the indicated want and belief, but also implies that this action-plan *caused* (in the characteristic way) the act of flipping the switch. (p. 78)

This doesn't say in so many words that the want and belief are functioning as antecedent causes of the action, rather than as contemporaneous causes; but the addition of the disjunct "resulted in" strongly suggests this. This impression is reinforced by the fact that when Goldman first speaks of want-belief causation of actions he authorizes a number of other terms that plainly imply a temporal sequence of cause and effect.

Instead of saying that wants and beliefs "cause" acts, one might say that acts "result from" wants and beliefs, or that they "stem from" wants and beliefs, or that they "flow from" wants and beliefs. (p. 55)

Goldman also discusses various issues in ways that make it clear that he is thinking of wants and beliefs as antecedent causes. For example, at one point he argues that a complex act like taking ten steps might be, as a

whole, a basic act, in which case each single step need not be caused by a separate want-belief pair; it would be sufficient for a single want-belief pair to cause the entire sequence. Note how, in making this point, he speaks of the causal relation.

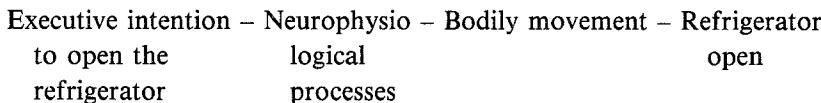
A desire to take ten steps might cause an act-token of taking ten steps without the necessity of having a different occurrent desire *immediately prior* to each of the ten steps . . . any temporal part of an act-token is an act-token . . . Then each of the ten steps taken during this interval is also an act-token, despite the fact that there was no distinct occurrent want *immediately preceding*, and directed at, each distinct step . . . This shows that not every tiny act-token is caused by an *immediately preceding* occurrent want to do *it*. In other words, an act-token which is part of a larger act-token must be caused by an occurrent want, but not necessarily by an occurrent want specific to and immediately preceding it. (p. 89; emphasis added)³

Consider too the way in which causal theorists worry about "wayward causal chains". This problem is most usually introduced in connection with a *want-belief causation* account of what makes an action intentional, rather than in connection with the CEI of PEA's; but it nonetheless reveals something about how causal theorists are thinking of want-belief causation of action. Here is a typical example. "Suppose a man believes that if he kills his uncle he will inherit a fortune and suppose he desires to inherit a fortune; this belief and desire may agitate him and cause him to drive in such a way that he accidentally kills his uncle."⁴ This is a counter-example to a simple want-belief causation theory of intentional action; for even though the action of killing his uncle was caused by an appropriate want-belief pair it is not intentional. It is also a counter-example to a simple CEI, since even though the action was caused by the want-belief pair in question, it is not the case that the man killed his uncle in order to inherit a fortune; the killing being accidental, it was not done for any purpose at all.

These difficulties involve some non-standard links in the causal chain leading from the want and belief to the action. (Hence the label "wayward causal chains".) The want-belief gives rise to the action by way of some emotional upset, or by way of some machinery external to the agent, or . . . Now my present point is simply that in taking such cases to be a difficulty, causal theorists like Goldman show themselves to suppose wants and belief to function as *antecedent* causes of actions. Otherwise there would be no (temporal) room for atypical intermediaries, and the cases would pose no problem for a causal account.

Just how are we to think of a Want-Belief (W-B) pair as causally contributing to the performance of an action? To give a full answer to this question we would have to lay out and defend an account of the nature of action; and we can't undertake that here. However it will be useful to have a model of the structure of an intentional action and of the way in which a W-B pair can causally contribute to its occurrence. I shall present what I take to be the most defensible way of thinking of this, though I shall not have the space to defend my choice, and I shall have to leave many issues unresolved. I also believe that this model fairly well represents the way many advocates of the CEI think of the matter, sometimes only inchoately.

Let's go back to opening the refrigerator. It is clear that necessary conditions for my opening the refrigerator include (a) some appropriate movement of my body (parts thereof) and (b) this resulting in the refrigerator's door being open. Moreover, it is reasonable to suppose that in order for it to be true that I intentionally opened the refrigerator it is not only necessary that (c) I had what we might call an "executive intention" (otherwise called a "volition") to do so, at least at the time of action, but that (d) this executive intention played a crucial causal role in bringing about the bodily movement referred to in (a).⁵ So this action involves at least the following structure.



This diagram may well be too simple to fit the facts. For example, it might be that the intention that proximately sets off the neurophysiological processes is an intention to make just those overt movements. In that case the intention to bring about the final goal (refrigerator being open) would give rise to the proximate intention to make just those overt movements. Or perhaps that latter intention is brought about by *wanting* the refrigerator to be open and *believing* that performing movements of those sorts will bring that about.

Now if wants and beliefs are antecedent causes of action we will have to determine when the action begins, so as to be able to fix the point before which these causes are to be located. And this is a matter of intense

controversy in the theory of action. Is the action to be identified with the entire sequence from executive intention to result aimed at, or is it to be identified with some portion thereof, and if the latter which portion? Virtually every conceivable choice has had its advocates. Among the parts, the action has been taken to be the volition ("executive intention"), the bodily movement, the causal sequence involving neurophysiological processes leading to bodily movement, and the latter plus the intended result. For our purposes it is not necessary to make a choice between these alternatives. It will suffice to note that no one pushes the action back beyond the activation of the executive intention. And in terms of our diagram the most plausible causal role for an appropriate want and belief (in this case a want for some beer and the belief that opening the refrigerator would be a good way to get some) is as what brings about that activation. As a result of wanting some beer (more than anything incompatible at the moment) and believing that opening the refrigerator would be the best (or a good) way of getting some, one forms the present intention to open the refrigerator. That is certainly a plausible story. And any causal operation of the want-belief at a later stage would not be a clear case of antecedent (to the action) causation. If the want and belief were, e.g., to causally contribute to the occurrence of the bodily movement, after the intention had set off the relevant neuro-physiological processes, that would presumably count as a causal influence *during* the action rather than *prior* to the action.⁶ Hence let's take it that an antecedent want-belief pair causes the action by way of activating the executive intention, which in turn initiates the process that leads to the overt bodily movements and their intended result.

Let's note in this connection that when an explanandum itself involves a causal sequence it is quite common for the citation of a cause of the initiation of that sequence to be called a causal explanation of the whole complex. Thus if I want a causal explanation of the fact that your house burned down (that the house disintegrated as a result of fire) I will normally be looking for an account of how the fire started. And if I ask you why you got fired (why your job was terminated as a result of the action of your employer) I will typically be looking for what led your employer to act in that way.

Another relevant issue is that between an event-causality and an agent-causality account of action. The above diagram embodies an event-

causality account. Let's take the agent-causality account to differ from this in the following respect. Instead of thinking of the executive intention as just "happening" in the agent, it thinks of *the agent's forming the intention*, where that fact is neither a matter of some event-causal transaction nor determined by event-causality, though it can, of course, be influenced by factors like wants and beliefs. On the agent-causality view, wants, beliefs, and other motivational factors "incline without necessitating". The agent takes them into account in forming his intention, but whatever they may be and however strongly they push in a certain direction, the agent has the power to guide his activity in some other direction. Thus intention formation represents a "break" in the flow of event causality. I am not able to go into this issue. My present point is simply that the substitution in our diagram of "The agent forms the intention to . . ." for "The intention to . . . occurs" will make no differences as to where the want-belief pair cited in the PEA makes its causal contribution. That contribution will still be to the formation of the intention. Of course the want-belief pair will no longer be thought of as forming part of a causally sufficient condition for the intention formation. But it still remains true that such causal contribution as it makes to the action comes from being an influence on the agent in the formation of the intention.

II

The CEI has been hotly contested in recent decades. Some of the objections I take to be have been definitively answered and I shall not discuss them here, e.g., the claim that wants and beliefs cannot cause actions since they are "logically connected" with actions, and the argument from the non-existence of laws connecting wants and beliefs with actions.⁷ However there are other objections that I take to be worthy of more respect. I am going to mention one of those objections now. I will not develop it as I would if I were aiming to refute the CEI at this point. Instead I shall present it as something that may properly worry us about the CEI, and thus motivate us to search for an alternative. After presenting the alternative I will develop this objection into an argument for the superiority of that alternative.

It has been urged against the CEI that, so far from a PEA being a

specification of antecedent causes, it is rather a *redescription* of the action. To explain my getting up early by my want to get some yard work done is to redescribe my getting up early as a preparation for doing yard work. To explain my signing a check by reference to my wanting to pay the water bill is to redescribe the check signing as paying the water bill. Thus stated, the objection is quite weak. Davidson effectively disposes of it by the simple observation that one can redescribe an event in terms of its causes.⁸ One can redescribe a door's opening as a case of a door's being blown open, i.e., redescribe it in terms of its being caused by a gust of wind. Therefore there need be no competition between the CEI and the redescription interpretation.

Nevertheless there may be blood in this turnip. Whatever the intentions of the philosophers who have deployed this argument, I see behind it a sense that there is something misguided in supposing that a PEA is concerned with the *causal history* of an action. When I say that *S* did *A* in order to *B* it doesn't feel to me as if I am delving into etiology; I don't have the sense that I am contributing to a case history. It seems rather that I am bringing out some factor or aspect of *A* itself, some facet of the matter that is strictly contemporary with *A*. It seems that I am making a claim about what *S*'s purpose is *in* doing *A*. And this sense of the matter is even stronger when I explain a current action of my own. In telling you that I am opening the window in order to get some fresh air it doesn't seem to me that I am telling you something about how that action was generated. It seems rather that I am telling you something about what I am "up to" here and now, what I am opening the window for (now), not what led up to my doing so. Of course these impressions may be misleading. But if we can find an otherwise satisfactory account that accommodates them, so much the better.

These considerations have commonly been used to support what I will call "radically non-causal" interpretations, interpretations according to which in putting forward a PEA I am not purporting to say anything about what is "responsible" for the fact that *S* is doing *A* at the moment rather than anything else. In giving a PEA I am not aspiring to throw light on the fact that the course of events has taken this turn. Such interpretations include the following.⁹

(10) A PEA exhibits the action as rational, justified, the thing to

do, from the agent's point of view, i.e., given the agent's goals, wants, needs, principles, beliefs, etc.

(11) A PEA makes explicit how the agent was thinking of the action, what he was doing it as. Some versions of this view are couched in terms of the "meaning" or "significance" of the action (to the agent). They assimilate PEA's to the interpretation of utterances or texts and hence are can be called "hermeneutic" interpretations.

These views have been definitively criticised by Hempel, Davidson, and others,¹⁰ and I will simply ride piggy-back on those criticisms. The basic point is that a PEA is designed to throw some light on the fact that *S* is doing *A* at *t*, rather than doing any of the innumerable other things he might be doing. And it seeks to do this by specifying part of what is "responsible" for that state of affairs, part of what "brings it about" that *S*'s doing *A* is what we have at that juncture. The radically non-causal views miss that crucial feature. I am not denying that it is possible, legitimate, important, or interesting, to seek explanations like these. Nor am I denying that they are properly called 'explanations' (though I do have reservations about that). I am merely pointing out that they lack some of the central intended force of PEA's.

III

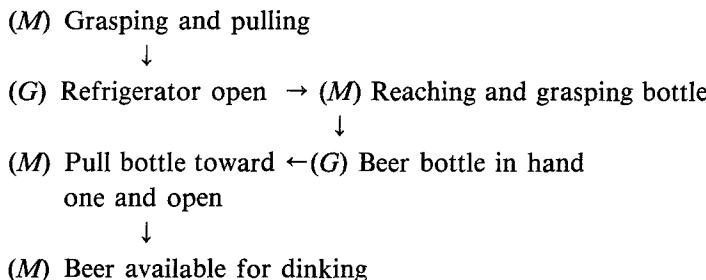
Where does *this* leave us? We found reasons to be dissatisfied with the CEI, but those reasons seem to push us, and have usually been taken to push us, in the direction of views that radically misconstrue the thrust of PEA's. And when we appreciate this point we seem to be forced back to a CEI. Is there any escape from this oscillation? Can we find an interpretation that will at once specify what is responsible for *S*'s *A*'ing and also avoid what we found to be objectionable in the CEI?

Yes, I believe that we can. The clue is found in the fact that our objection to the CEI centered around the point that the CEI represents a PEA as citing *antecedent causes* for the action. This suggests that we could avoid the difficulty, while retaining the "why it happened" thrust, if we could construe the PEA as identifying some crucial *contemporary* influence on the action. I will now proceed to follow out this suggestion.

Thus far we have been following the usual practice in action-theory literature (or at least the impression usually given by that literature) in thinking of the motivation of action as involving a causal sequence of events, each of which ceases, or fades into the background, when its successor(s) arrives. An activated want and belief give rise to an intention, and then cease to be, at least in an activated form. Their influence was restricted to the initiation of the intention. Again, the intention gives rise to the bodily activity (via some neural mechanism), and then it can fade from the scene, having played its appointed part in the drama. But this is false to the actual facts of human intentional action. In particular, an executive intention does not simply push the button to set off the appropriate neuro-physiological mechanism and then retire to the wings. On the contrary, there is a mental direction, monitoring, and guidance throughout the duration of the activity. This is most obvious for actions that take an extended period of time. If the action is writing a paper, baking a cake, or mowing a lawn, we can hardly suppose that no mental, intentional factors exercises any influence after the initiation of the activity. The carrying out of the intention is *not* just a matter of the unrolling of automatic, pre-programmed physiological sequences. On the contrary, one is, often consciously, monitoring the progress of the activity, making corrections when one goes off course, initiating sub-actions, making decisions as to what sub-routine to activate, and so on. And though this is most obvious for complex actions it is not restricted to those. Even so simple and routinised an activity as opening a refrigerator will exhibit this structure. To be sure, after one has thoroughly learned how to do this, one will not ordinarily be conscious of any higher mental direction. But if something goes wrong one is immediately aware of it, and one finds oneself figuring out how to remedy it. If the handle is stuck or breaks off I will immediately be aware of that, and I will try to devise some other way of getting the door open. If my fingers won't move, or if I fail to grasp the door tightly enough, I will become aware of that and try to correct the situation. Even in such a simple, overlearned action as this, a representation of the goal and the chosen path to the goal is activated throughout the activity, and the bodily activity and its results are being monitored in the light of that representation.

Let's call the complex representation that is employed in the monitoring and guidance of ongoing behavior an "action-plan".¹¹ At a minimum an

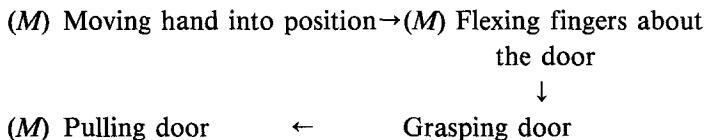
action plan will consist only of *the representation of* (I shall normally omit this qualification, leaving it to be understood) a single goal state; this will be the case, presumably, with the simplest basic actions, where the role of intentional psychological factors is merely to "specify" the type of bodily movement to be performed; everything else is left up to automatic physiological mechanisms.¹² Above this level action plans will be more complex. There will at least be a goal and some more or less complicated means to that goal, involving at least some bodily movements, and also, in many cases, preliminary goals and means for achieving them. Thus an action plan that guides getting a bottle of beer may include the action plan that guides opening the refrigerator as a part. This larger plan will look something like this. ('M' for means and 'G' for goal.)



Of course, what figures in the action plan as a goal could also be conceptualized in action terms. For example, the second goal above could just as well be specified as *holding the beer bottle*.

This diagram may not adequately represent the full complexity of this particular action plan. For example, the plan may break the initial movement (grasping and pulling) into components, each of which would put the agent in a position to carry out the next stage. At the least we would probably get a distinction between grasping the handle and pulling it. And conceivably the grasping would be broken down into (a) moving the hand into position and (b) flexing the fingers around the handle. I am not now speaking about the actual movements; of course they have these distinguishable phases. I am speaking of the mental representations that are used in the guidance of behavior. Clearly not every distinguishable feature of the bodily processes involved will be presented in the action plan. We can be confident, e.g., that neurophysiological details of which

the average agent knows nothing will not be included. In just how fine a detail the peripheral bodily movements are represented in the action plan is a detailed question for psychology. In any event, let's take it that the first line of the above diagram should really include the following sequence of movements, each of which puts the agent into a position to perform the next.



Let me say another word about the issues involved in determining how far down actions plans extend. It may well be that as behavioral routines become thoroughly learned more is left to "automatic" mechanisms, and less is represented in plans. It is certainly true that mental representations play a larger conscious role in action guidance when one is first learning how to do something. But we should not conclude from this alone that mental representations are playing no role in routinized actions. Action plans can function unconsciously. Normally I am not conscious of the various particular movements I make in typing or driving a car, so long as things are going smoothly; but when a hitch develops I am immediately aware of it and seek to take corrective action. It is plausible to explain this by postulating an unconscious monitoring of the behavior in the light of a representation of how things are supposed to go. But the final word of just when such postulations are warranted belongs to psychological theory.

Of course, an action plan, i.e., a structured representation of goals and means thereto that is suitable for use in the direction of current behavior, is not always so used when psychologically activated. Such a plan is typically before the mind in deliberation as to what one is to do. Here, however, we are specifically interested in the use of plans in the direction of activity. An action plan figures in the control of ongoing activity by providing a "schema" or "template" with which the developing activity and its results can be compared. In the initial stage of opening the refrigerator the action plan provides a representation of the desired goal (refrigerator open) and the chosen means thereto (grasping the handle of

the refrigerator door and pulling outward). So long as the flow of sensory information "matches" that template all is well. But whenever the feedback is negative the agent will seek to modify the activity until it is back on course, i.e., until what is actually happening matches the template.

Many issues would have to be gone into before the concept of an action-plan is sufficiently determinate for theoretical employment. Most of that task lies outside the bounds of this paper. But we must tie down one loose end, since it is crucial for the action-plan interpretation (API) of PEA's that is in the offing. How are we to understand the arrows in our diagrams? That is, how is the *in-order-to* relationship represented in an AP? We don't want the final word on this to be that the AP uses the concept of *in-order-to* in its representation. For we want to elucidate *in-order-to* explanations by reference to the role of AP's in the guidance of behavior. And to find that very concept, in undigested form, in the explanans would be disappointing, to say the least. But what alternative have we?

It is tempting to answer this question in terms of means-end beliefs. M_1 is represented as done in order to achieve G_1 by virtue of the fact that S believes that M_1 will (have a good chance to) lead to G_1 . Of course, as noted earlier, there is a variety of such means-end beliefs. Sometimes it is believed that M will cause, or causally contribute to, G , sometimes that M will put one in a position to bring about G , sometimes that M will, in the circumstances, constitue reaching G , and so on. But let's assume that we can either find a satisfactory common formulation, perhaps in terms of M 's (probably) *giving rise to* G , or else can formulate an exhaustive disjunction of the variant forms. Even so reference to such beliefs will not do the job. A minor difficulty is that we have not taken AP's to include beliefs. But we gave no reason for that exclusion, and the present consideration might well outweigh any such reasons. And, in any event, the structuring of AP's could reflect S 's beliefs, even if they are not "in" the AP. A more important consideration is this. One may have many means-ends beliefs that do not give rise to AP structuring. For example I may believe that opening the refrigerator will displace air molecules, as well as put me in a position to take out a bottle of beer. And yet the former belief makes no contribution to the AP in question. It may be replied that the displacement of air molecules is not even in the

AP in question; so that we don't have to explain why it is not being represented *as* an end *vis-à-vis* opening the refrigerator as a means. True enough; but that just pushes us back to the question of how to determine what is and is not included in a given AP. It is not as if we can segregate the brain into distinct AP's by anatomical dissection.

I take it that these considerations are pushing us in the direction of an account in terms of the function of AP's in the monitoring and guidance of behavior. The representations that make up an "activated" AP are those that are used in the monitoring and guidance of behavior. They are representations of what the control system looks for to decide whether things are satisfactory and, if not, what to do about it. No representation of air-molecule-displacement is in this AP, for nothing in the way of behavior-direction depends on whether air molecules are being displaced, or if so how many and in what manner. Whereas the control system is sensitive to whether the hand is in a certain position, whether the fingers are sufficiently firmly grasping the handle, and so on. As we put it above, the representations in an AP constitute a "template" with which the ongoing course of behavior is matched. When a mismatch appears, corrective steps are undertaken.

But even though this may answer the question: "What are the entrance requirements for AP's?", it does not yet give us an interpretation of the arrows; i.e., it does not yet tell us how the means-end relationship figures in the AP. But to achieve this we only need to extend the above point. Both *M*'s and *G*'s function in the template, but it is the latter that play the dominant role. The dominance manifests itself in two ways. First, it is only because it is believed to have a good chance of giving rise to the *G* that the *M* is there in the first place. Were it not for its presumed role in getting the refrigerator open there would be no interest in monitoring behavior for degree of match to the *grasping and pulling* template.¹³ (Lest it be thought that we have regressed to the earlier criticized view that means-ends beliefs carry the load, let me point out that the present suggestion is not that the mere existence of the belief does it. It is rather that *M* and *G* are distinguished by virtue of the fact that *M* plays a role in the monitoring and guidance of behavior only because it is believed to have a good chance of contributing to the *G*; and not vice versa.) Second, the dominance is reflected in what happens after admission, as well as in the criteria of admission. Behaviour will be held to conformity

to the *M* template only with respect to its efficacy in bringing about the *G*. If, e.g., the door should swing open of its own accord, the control system would lose its interest in the grasping and pulling motions and pass on to the next stage. In other words, *M* is retained, as well as admitted, only so long as it is taken to have the right relationship to *G*; but not vice versa.

Here are a couple of further points about AP's.

1. Action plans are not perfectly determinate. For one thing they can hardly make provision for all contingencies. It is true that they can involve branchings: if the door is locked open it with this key; if not just grasp the handle and pull outward. But no matter how many such alternatives are included, it will always be possible that things do not go as anticipated. In that case improvisation is called for. Moreover even where things go according to plan not every detail is represented in the plan. Just exactly how far is one to open the refrigerator door in order to get the beer? The typical plan for this activity will leave that open.

2. Both event causality and agent causality approaches to action can recognize the role of action plans in the guidance of behavior. They will differ on what is involved in the putting into effect of action plans, and on the conditions under which plans can be changed in mid-course. More specifically, an agent causality theory will insist that the activation of, and changes in, an action plan will, normally, be a matter of agent causality, while the event causality theorist will take all this as due to event causality. Nevertheless, they need not differ as to the role action plans play in the monitoring and guidance of behavior.

IV

Once we grasp the point that ongoing intentional behavior is under the control of action plans, we see the possibility of an interpretation of PEA's that combines the crucial desiderata. In explaining an action by citing its purpose we are making explicit part of the structure of the action plan in terms of which that action is being monitored and guided. More specifically, we are making explicit the goal that is responsible for the presence and functioning of the explanandum in the action plan, the goal on which that action's inclusion in the plan is functionally dependent. To wit.

- (12) S opened the refrigerator at t in order to get a bottle of beer.
– *Opening the refrigerator* was in the action plan that was guiding S 's action of opening the refrigerator at t because S believed at t that opening the refrigerator would enable him to get a bottle of beer.
- (13) S gave Smith the check in order to repay a debt. – *Giving Smith a check* was in the action plan that was guiding S 's action of giving Smith a check at t because S believed at t that giving Smith a check, in those circumstances, would constitute repaying a debt.
- (4) S flipped the switch at t in order to turn on the light. – *Flipping the switch* was in the action plan that was guiding S 's action of flipping the switch at t because S believed at t that flipping the switch would cause the light to go on.

And as a limiting case we have “He took a walk just because he felt like it”. That is to say that taking a walk is in the current action plan on its own; it is not functionally dependent on anything else in the plan.

Let's make explicit how this action plan interpretation (API) accommodates the various requirements we have mentioned.

1. It is obviously in harmony with the sense that in proffering a PEA we are bringing out features of what is going on now, rather than giving an etiology. The action plan is operative while the action to be explained is in progress.¹⁴

2. The API makes intelligible not only the fact that a person seems to have immediate and privileged epistemic access to the purposes for which he is doing what he is doing, but also the fact that one's awareness of *why* he is doing *A* seems to be part and parcel of one's awareness *that* one is doing *A*. The *why* seems to come with the *what* in the same package. On the API this is all quite understandable because both my awareness of what I am doing and my awareness of why I am doing it consist in awarenesses of features of the currently operative action plan. They have the same source and can be expected to have the same epistemic status and to “feel” as if they come together. In making this point I am not taking agents to be omniscient and infallible either about what they are doing or why they are doing it. There is room for error. I hold that by and large people do have immediate, privileged knowledge of these

matters; but this assumption is compatible with the recognition that such knowledge is both fallible and incomplete.

3. On the API a PEA does specify something that is (partly) responsible for the fact that *S* is doing *A* now rather than something else. For *S* is *A*'ing because *A*'ing is in the current action plan; if *A*'ing weren't included in the action plan *S* would not be *A*'ing, at least not intentionally. And *A*'ing is included because it is believed to be related in the specified way to the goal, *G*, that is cited in the PEA. Thus the PEA does identify an explanans on which the explanandum is functionally dependent. It doesn't stand apart from the determination of the actual course of behavior.

But if the API takes a PEA to be specifying something (the AP or certain features thereof) that is (at least partly) responsible for the fact that *S* is opening the refrigerator, how can it be a wholly non-causal interpretation? Let's take it as established that the API is distinct from any interpretation in terms of *antecedent* causes, as the CEI has usually been thought of. But that leaves open the possibility that the API takes a PEA to do its explanatory job by specifying contemporaneous causal influences. More specifically, according to the API, in saying that *S* went to the office to clear up some correspondence I am saying that *S*'s going to the office when he did was partly due to the presence in the currently operative AP of a representation of going to the office, which in turn was there because of its believed relationship to getting correspondence done. Am I not thereby identifying a contemporaneous causal contributor to the action of going to the office? And is not this identification playing a crucial role in my explanatory enterprise?

Well, yes and no. The API does undoubtedly presuppose the existence of a certain causal contribution to the explanandum.¹⁵ It is not quite correct to say that it presupposes that the AP is a cause; for an AP is a complex mental representation and, as such, is not fitted to play a causal role. However there is no doubt but that an AP's *being in effect*, the *use* of the AP by the control system to guide and monitor behavior, is making a causal contribution. For the fact that *S* is doing *A* at the moment is functionally dependent on that AP, rather than any one of innumerable others, being the one that is used to monitor and guide behavior at that moment.¹⁶ According to the API what a PEA does is to make explicit part of the structure of a certain AP. What AP? The one that is being

used in behavior guidance with respect to that stretch of behavior that includes the action to be explained. Unless the PEA were presupposing that the AP, a part of whose structure it specifies, were playing that causal role, it would lack the intended explanatory force. But the point is that this is a *presupposition*. It is no part of what is being explicitly claimed by a PEA, on the API. On this construal a PEA does not consist in citing or specifying causes of the explanandum, claiming that they are causes. In this, as well as in temporal location of causes involved, it is distinguished from the usual CEI.

But this may well be considered a superficial difference. After all, whether I explicitly assert that there is exactly one king of France at present, or rather presuppose this, in either case I *commit* myself to this being the case in what I say. Similarly, even if I am presupposing, rather than explicitly asserting, that the activation and use of a certain AP is a causal influence on the explanandum, so long as the force of my explanation depends on the truth of that assumption I am committed to that assumption as firmly as if I had flatly asserted it. And isn't that sufficient to render the PEA, on this construal, a causal interpretation?

Maybe. But there is still another difference from the standard CEI. Even the presupposition is much less specific than the claim that an antecedent want and belief are among the causes of the action. More precisely stated, the presupposition is that *S* is doing *A* at *t* (in part) because an action plan involving a representation of *A* is being used by *S* (or *S*'s control center) in the guidance of *S*'s behavior at *t*. This includes the claim that there is some kind of causal mechanism in operation at *t* that involves a flow of information concerning *S*'s overt behavior and the supervision of the details of that behavior in the light of that information. But the nature of this mechanism is not specified. No parts of it are mentioned; its general character is left open. Does it involve neural processes and if so, what sorts? Does it involve the operation of concepts, imaginative schemata, or reasoning? Are activated wants and beliefs playing any causal role? All this is left up in the air. The presupposition is simply that somehow the AP in question is used in a monitoring and guidance function; and that the actual performance of the action being explained is somehow functionally dependent on that.

A computer analogy may help. An action plan is like the program, or

portion of a program, on which the computer is currently operating; and the putting an action plan into effect is like activating a program, or portion thereof. Now suppose that my word processing program involves changing the placement of, e.g., a paragraph, by first making a space and then inserting the paragraph there. Then in saying that the computer is making a space in order to insert a certain paragraph therein I am making explicit a certain feature of the structure of the program that is currently activated. No doubt the execution of the program requires certain appropriate causal processes in the "hardware", but I am not saying anything about what that is like. I make a maximally unspecific presupposition of an underlying causal structure in supposing the program to be activated, but what I explicitly cite in giving my explanation is the structure of the program activated.

Thus the API differs from the standard CEI not only in not representing a PEA as citing antecedent causes, but also in not representing it as being committed to any specific causes at all, even by way of a presupposition. If, given all that, one still wants to regard the API as a causal interpretation, he is at liberty to do so. But the fact remains that it is a quite different sort of interpretation from the standard CEI.

V

Now I should like to provide some additional arguments for preferring the API to the usual CEI in terms of antecedent causes. These arguments are designed to support the claim that our intent in giving PEA's is to bring out current structure rather than causal history.

First, consider the bearing on PEA's of drastic deviations from normal etiology. Suppose that *S* has an idiosyncratic psychological organization such that, from time to time, action plans emerge fully formed "from nowhere", i.e., without resulting from normal processes of intention formation. *S*'s psyche is such that he stores various action plans, which, from time to time, emerge with no discernible relation to current motivational processes.¹⁷ Let's suppose that this is true of our refrigerator opening. *S*'s current intention was not generated by an antecedent want for some beer and a belief that opening the refrigerator is the way to get it. On the usual CEI it would be false that he is opening the refrigerator

in order to get some beer. But couldn't that be true, nonetheless? So long as his current action plan is as we have been supposing, that "in order to" account is correct, whatever the causal history.

A variation on this theme. Suppose we are fundamentally mistaken about what is normal in action plan formation. Suppose the activated wants and beliefs we are aware of in deliberation, practical reasoning, and less explicit versions thereof, are mere epiphenomena, thrown off by more fundamental psycho-physiological processes of which we know nothing. Even so, if the action plan that is currently controlling *S*'s behavior is as we have specified, the PEA would be accurate. It would still be true that *S* is opening the refrigerator in order to get some beer.

Third, consider last minute changes in plans. Suppose that the intention to open the refrigerator was formed because of a desire for some gin and tonic and a belief that tonic is to be had in the refrigerator. But then just as I am beginning to open the door, it occurs to me that a cold beer would be much more gratifying, and I continue my activity but with *getting some beer* substituted in the controlling action plan for *getting some tonic*. From that moment on it would be true that I was opening the refrigerator to get some beer, even though on the usual (antecedent) causal interpretation it would be true that I was opening the refrigerator to get some tonic. Epistemologists will recognize here an analogue of the point that whether my belief is justified depends on what is currently sustaining it rather than on what originally engendered it, where these deviate. If I came to believe something from trusting idle gossip, but later believe it because I have good reasons for it, my belief is justified at that later time because of what sustains it then.

Finally, note that the API relieves us of the problem of "wayward causal chains". Remember that, as applied to the CEI of PEA's, the difficulty is that a want-belief pair may cause an action without thereby grounding a correlated PEA, if the causal sequence is of the wrong sort. I may desire to get the attention of the speaker and believe that I can do so by raising my arm. This want and belief may lead to an agitated condition, which in turn results in my arm going up. But it is still not true that I raised my arm *in order to* attract the attention of the speaker, or *in order to* anything else. The arm raising was not purposive at all. This leads causal theorists like Goldman to build in the requirement that the want and belief cause the action "in the characteristic way";¹⁸ but they

are not able to provide an adequate specification of this "characteristic way". The API is not faced with these difficulties, and for two reasons. First, it does not represent the PEA as providing the explanation by citing causes. Hence it is not vulnerable to the possibility that the specified causes might bring about the action in some non-purposive fashion. And, second, even if it did specify causes they would not be *antecedent* causes; and so there would be no room to insert wayward chains between the causes and the action.

Let me emphasize the point that I am not denying (the API does not deny) that actions have antecedent causes of the sort envisaged in the CEI. (Nor am I asserting that they have such causes.) My position, rather, is that whatever the causal history of actions may be, the PEA's we actually give, as we normally understand them, do not specify such causes. Thus my opposition to the CEI is much less radical than the usual objections that are based on the denial that actions can be caused or that they can be caused by wants and beliefs. I am not denying that explanations along the line of the CEI could be given; for all I know, they are sometimes given. Explanation depends on what we suppose we already know or understand, and what specific gaps we are trying to fill. Just as one might well seek the kind of understanding given by PEA's on the hermeneutic interpretation, so one might well seek to identify the wants and beliefs that played a crucial role in the formation of an intention or a derivative want. My contention is only that the CEI misconstrues what we are normally up to in giving PEA's.

VI

Finally I would like to say something about the relation of the API to a nomological interpretation of PEA's. This requires separate treatment, since its advocates often disavow any commitment to the view that PEA's specify *causes* of actions.¹⁹ The general idea is that appeal to wants and beliefs in PEA's has explanatory force because of the existence of laws, at least probabilistic or tendency laws, in which wants and beliefs figure as conditions of actions. Thus on this view PEA's are of generically the same sort as explanations of nonpersonal phenomena that are based on laws of nature. We can explain the window's breaking by saying that a large rock hit it traveling at a considerable speed, because there are laws

that specify impact with a certain degree of force as a sufficient condition for shattering of glass that has certain physical properties. Similarly, it is supposed we can explain an action by citing wants and beliefs because there are laws that represent actions as determined by wants, beliefs, and the like.

Can we actually formulate laws that connect wants and beliefs to actions in the required way and are plausibly regarded as true? Well, we can't come up with anything very tight, but there are approximations. Here are a couple of examples.

- (15) If x wants p then, under favorable conditions, if x judges that doing A will probably lead to p and that not doing A will probably lead to not- p , x will feel some impulse to do A .²⁰
- (16) If x wants p , then for any action or activity A which x has the ability and the opportunity to perform, if x believes either (a) that his doing A is necessary to p , or (b) that his doing A would have at least some considerable probability of leading to p , or (c) that his doing A would have at least some considerable probability of constituting attainment of p , then x has a tendency to do A .²¹

The consequent in these formulations is not the action itself but rather an impulse or a tendency. These could be called "tendency laws". Whether the tendency in a particular instance actually issues in doing A depends on various other factors: the abilities and opportunities of the agent, and the relative strength of competing contemporary tendencies.

Now if (15) and (16) are to be understood as relating *antecedent* wants and beliefs to actions, or to action tendencies, then the API will differ from the nomological interpretation in the same way it differ from the standard CEI; and the same reasons for the superiority of the API will be applicable. And it does seem that these formulations are naturally interpreted in that way. The thought behind them would seem to be that a suitable want-belief pair, a want for a goal and belief as to what action will enable one to get there, will *give rise to* a tendency to perform that action, and that what tendency *issues* in action depends on the sorts of factors mentioned in the last paragraph. This is why (15) and (16) have an action tendency in the consequent. In other words, these would seem

to be laws that govern (if anything) the formation of executive intentions or action plans.

But a nomological theorist could suppose that the laws in question relate contemporaneous wants and beliefs to actions, or at least action tendencies. Indeed, Robert Audi has assured me that this is the way he was thinking of the matter when he formulated (16). In that case the API will lose the edge it possesses by virtue of dealing with factors contemporary with the action. If it is superior to this contemporaneity form of the nomological interpretation that will be because it gets contemporary explanatory factors into the picture in the right way. And such I believe to be the case. The API will differ from all forms of the nomological interpretation in that it does not presuppose any laws linking actions with wants and beliefs, or with anything else. The assumption of such laws is no part of the content or explanatory force of a PEA, according to the API. That is not to say that it denies that actions figure in the consequents of some (deterministic or probabilistic) laws. It simply makes no commitment to anything of the sort. Once again the API is noncommittal *vis-à-vis* its rivals. And once again this turns out to be a virtue. It seems to me quite implausible to suppose that whenever anyone proffers a purposive explanation he is assuming any such laws as those envisaged by nomological theorists. Again, he may be presupposing that there are appropriate causal processes involved in the operation of action plans; but that does not carry with it any assumption of laws connecting actions with wants and beliefs, or with any other particular explanatory factors. Once again, the API wins by attempting less. Small is beautiful.

NOTES

¹ This is not to deny that other information is provided as well. Thus (6) specifies an obligation as involved, while (5) represents the agent as having this aim because he likes doing yard work.

² See Davidson (1963), Goldman (1970, esp. Ch. 3, Section 6), and Tuomela (1977, esp. Ch. 8).

³ Davidson also lets it slip, in passing, that he is thinking of antecedent causes. "Mention of a causal condition for an event gives a cause only on the assumption that there was also a preceding event. But what is the preceding event that causes an action? In many cases it is not difficult at all to find events very closely associated with the primary reason. States and dispositions are not events, but the onslaught of a state or disposition is. A desire to

hurt your feelings may spring up the moment you anger me; I may start wanting to eat a melon just when I see one; and beliefs may begin at the moment we notice, perceive, learn or remember something." (p. 694)

It is also relevant to note that those who put forward causal theories of the nature of action in terms of want-belief causation of action also typically think of wants and beliefs as antecedent causes of action. See, e.g., Brand (1984, pp. 6, 16, 17, 20, 31, 34, 35).

⁴ Chisholm (1964, p. 616). Cited and discussed in Goldman (1970, pp. 55 ff.)

⁵ We may think of the "adoption" of the "executive intention" as the psychological occurrence that triggers the neural processes that in turn trigger the peripheral bodily movements involved in the action. The crucial role of "present intentions" in the structure of action has recently been emphasized by Searle (1983) and Brand (1984).

⁶ The alternative interpretation we will develop involves something like a contemporaneous causal influence, though not precisely from wants and beliefs.

⁷ See Davidson (1963); Alston (1966).

⁸ 1963, p. 692.

⁹ For examples of such views see Mischel (1963); Dray (1957, ch. V).

¹⁰ See Davidson (1963); Hempel (1962).

¹¹ We might think of an action plan (one that is used in the guidance of current behavior) as just what we were earlier calling an "executive intention", one that is continuing its function into the period of the action intended. The term 'action plan', and similar terms, have been used by various other theorists. See e.g., Goldman (1970, Ch. 3, Section 3), Tuomela (1977, Ch. 7), and Brand (1984, Ch. 1 and Pt. IV). My concept differs somewhat from all of these, though there are certainly affinities. Thus Goldman takes an action plan to be a combination of a structure of beliefs and action-wants (pp. 23 ff.). Neither Goldman nor Brand suggest an interpretation of PEA's in terms of action plans. Tuomela may do so, though the details of his account of such explanations are not clear to me.

¹² It is a detailed question of psychology at what level(s) this is the case. This minimal description may not hold for everything that philosophers typically count as basic actions. For example, the bodily movements involved in grasping and lifting a hammer may normally be guided by mental representations of various sub-movements of fingers and arm. But it seems clear, from infinite regress considerations if from no others, that at some level the carrying out of an intention to do *A* is not guided by either a representation of some means to doing *A* or by a representation of some portion of *A*.

¹³ This statement needs complicating in various ways. For one thing, even if I weren't interested in getting the refrigerator open I might have other reasons for grasping the door, perhaps just this way. Indeed, I might have both these interests conjointly, in which case the current AP is really more complex than we have been supposing. Secondly, the grasping and pulling motions might be of interest in their own right as well as in their role as means to getting the refrigerator open. What this all boils down to is that a particular action plan will reflect all the relevant pro-attitudes and means-ends beliefs of the subject (at least all those that are currently activated); and this may well be more complex than a superficial assessment may suggest. The point is that if the earlier diagram accurately depicts the situation, the statement in the text is adequate.

¹⁴ It will not have escaped the reader's notice that in taking a PEA to appeal to an explanans that is contemporaneous with the action to be explained, the API makes a certain presupposition about the time of an overt action, viz., that it is being performed when the bodily movement involved is occurring. Since, as we noted earlier, there is a plurality of competing views as to what constitutes *the action*, this presupposition may be contested. However the

possibility of disagreement is minimized by the fact that the API need not assume that the action is exhausted, in temporal extent, by the duration of the bodily movement, but only that the action is in progress at least while the bodily movement is occurring. Even if the action extends before and/or after this in time, it will still be the case that the action plan is operative while the action is being performed. And the assumption that an overt action is being performed at least while the relevant bodily movement is occurring is virtually beyond controversy.

¹⁵ At least I am granting this for purposes of the discussion, Jonathan Bennett has pointed out to me that serious questions could be raised concerning the claim that everyone in every culture makes such a causal assumption whenever a PEA is proffered.

¹⁶ To be sure, one might dispute the claim that this dependence is properly called 'causal'; but this would undoubtedly be based on some of the discredited arguments against the standard CEI, to the effect that the likes of wants and beliefs cannot be causes of behavior.

¹⁷ A more drastic version of this would have it that *S* just came into existence complete with an operative action plan; but we need not go that far to make our point.

¹⁸ Goldman (1970, p. 62).

¹⁹ See Brandt and Kim (1963, pp. 434-435); Audi, (1973, p. 18), (1979, pp. 234-244).

²⁰ Brandt and Kim (1963, p. 427).

²¹ Audi (1973, p. 4). Audi presents this as a biconditional. For my purposes I have weakened it to a conditional.

REFERENCES

Alston, W. P.: 1966 'Wants, Actions, and Causal Explanations', in H. N. Castañeda (ed.), *Intentionality, Minds, and Perception*, Detroit: Wayne State University Press.

Audi, R.: 1973, 'The Concept of Wanting', *Philos. Stud.* **24**.

Audi, R.: 1979, 'Wants and Intentions in the Explanation of Action', *Journ. Theory Soc. Behavior* **9**.

Brand, M.: 1984, *Intending and Acting*, Cambridge, MA: MIT Press.

Brandt, R., and J. Kim: 1963, 'Wants as Explanations of Actions', *Journ. Philos.* **60**.

Chisholm, R. M.: 1964, 'The Descriptive Element in the Concept of Action', *Journ. Philos.* **61**.

Davidson, D.: 1963, 'Actions, Reasons, and Causes', *Journ. Philos.* **60**.

Dray, W.: 1957, *Laws and Explanation in History*, Oxford: Clarendon Press.

Goldman, A. I.: 1970, *A Theory of Human Action*, Englewood Cliffs, NJ: Prentice-Hall, Inc.

Hempel, C. G.: 1962, 'Rational Action', *Proc. and Addresses of the Amer. Philos. Assoc.* **35**.

Mischel, T.: 1963, 'Psychology and Explanations of Human Behavior', *Philos. Phen. Res.* **23**.

Searle, J.: 1983, *Intentionality*, Cambridge: Cambridge Univ. Press.

Tuomela, R.: 1977, *Human Action and Its Explanation*, Dordrecht: D. Reidel Pub. Co.

*Syracuse University,
Department of Philosophy,
Syracuse, NY 13210,
U.S.A.*