

Farewell To Davidson's Theory of Action

JOSEPH MARGOLIS

ABSTRACT

Two principal strategies for the Theory of Action are examined, Donald Davidson's and Georg Henrik von Wright's. Davidson's is shown to be entirely untenable, both in its own terms and in terms of the needs of the field; and von Wright's is shown to restore an option, opposed by Davidson, that saves a certain minimal concern of the field. The central issues turn on the inadequacy of the causal theory of action as formulated by Davidson and the quite independent matter that, under certain circumstances, the causal disjunction of reasons and actions is not convincing. The topics explored include: token physicalism, an extensional treatment of causality, anomalous monism, intentionality, reasons and causes, and the explanation of actions.

I

A recent overview of analytic theories of action confirms that the prevailing options have not changed substantially in at least the past twenty-five years¹. They appear to be polarized largely in favor of the arguments tendered by Donald Davidson in «Actions, Reasons, and Causes» (1963) and, substantially updated and refined, in «Mental Events» (1970)²; and in the rather opposed view of Georg Henrik von Wright, in *Explanation*

1 See Frederick Stoutland, «Philosophy of Action: Davidson, von Wright, and the Debate over Causation», in Guttorm Fløistad (ed.), *Contemporary Philosophy. A New Survey* (The Hague: Martinus Nijhoff, 1982).

2 Both essays are collected in Donald Davidson, *Essays on Actions and Events* (Oxford: Clarendon, 1980). Page references are made to this edition.

and Understanding (1971), which, in its own turn, updates a rather large literature that Davidson originally «defeated» in the earlier of the two papers mentioned – a series of books (and articles) that appeared, notably in the Monographs collected as *Studies in Philosophical Psychology* edited by R. R. Holland (Routledge and Kegan Paul) spanning the late fifties and the middle seventies³.

The truth is that Davidson's defeating arguments and alternative program are an unmitigated disaster and that von Wright's conterargument is essentially a restoration of the most minimal element featured in the pertinent contests over the analysis of action that, as it happens, Davidson sought to eliminate. Frederick Stoutland, who offers the survey, suggests that both Davidson and von Wright show some inclination toward reconciling their differences (perhaps Davidson needed to, more than von Wright) and achieving a rapprochement with the leading Continental European views. But Stoutland's attractive thought is certainly more wish than evidence. In any event, analytic theories of action tend to exhibit a kind of spurious rigor in the interval mentioned, largely in following Davidson's lead – not so much because they are formally defective but because adherence to their (admirable) rigor has misled an entire generation of younger scholars into thinking that what was being debated was very nearly all there was to the issue at stake. That impression, of course, needs to be explicitly combatted. It has lasted so terribly long.

Consider Davidson's theory. In the «Reasons, Causes, and Actions» paper, he advances what amounts to two essential claims: first, that causality and (numerical) identity are thoroughly extensional notions, but (also) that mental states and events («pro attitudes»: «desires, wantings, urges, promptings» and the like; plus believing: «knowing, perceiving, noticing, remembering» and the like) constitute what, following Davidson, we may call «*the primary reason*» for a person's performing an action – which both characterizes what persons do *as* actions and causally explains, by providing «justifying» reason, actions thereby performed⁴; and secondly, that causal explanation is a thoroughly intensional notion since it requires favorable descriptions for successful explanation (on which, in principle, rationalization depends) but also that rationalization (giving «a primary reason for an action») is, at bottom, a species of causal explanation⁵.

The point of Davidson's argument – the point of the stronger «Mental Events» paper as well – is to bring the general methodology and metaphysics of action into accord with allegedly canonical views about what a science is (in particular, what a scientific explanation is). It would not be unfair to say that Davidson favors Carl Hempel's well-known conception of the sciences, and inclines, however informally, toward a very strong (but distinctive) physicalism and an adherence to the unity of

3 See Georg Henrik Von Wright, *Explanation and Understanding* (Ithaca: Cornell University Press, 1971).

4 «Reasons, Causes, and Actions», pp. 3-4.

5 *Ibid.*, p. 12.

science program. In *these* respects, regarding the analysis of action and regarding the analysis of science, Davidson is either flatly mistaken or has not supplied adequate grounds for his claims, or has not answered obvious objections. The reason for contesting his views at this rather late date is simply that they seem to have sustained their magical influence almost unchecked, by some sort of extraordinary philosophical inertia. Since there are independent reasons for encouraging a rapprochement with «Continental» themes, it would be most helpful to be able to place the strongest «analytic» currents correctly so that they need not deflect us in this larger regard. Our objections should perhaps be cast as carefully as possible in an idiom congenial to the notable rigor of analytic philosophy. (It is almost impossible to exaggerate the influence of Davidson's model).

The argument of Davidson's «Mental Events» paper is, of course, the one that introduces and supports his well-known thesis that goes by the name of «anomalous monism»: in effect, the triad that includes «the Principle of Causal Interaction» («that at least some mental events interact causally with physical events»), «the Principle of Causal Interaction» («that at least some mental events interact causally with physical events»), «the Principle of the Nomological Character of Causality» («that where there is causality, there must be a law: events related as cause and effect fall under strict deterministic laws»); and the principle of «The Anomalism of the Mental» («that there are no strict deterministic laws on the basis of which mental events can be predicted and explained») ⁶. Davidson's primary concern here is to show that the triad is consistent and, therefore, promises to being the methodology and metaphysics of our treatment of action within the pale of the unity of science.

The essential counterclaims that need to be considered are these: (i) the triad that forms the doctrine of anomalous monism is inconsistent, essentially and irrecoverably; (ii) causality is not uniformly extensional; (iii) causality does not entail nomologicality; (iv) rationalization is not a species of causal explanation; (v) reasons, though they may be causes, need not be causes in contexts in which they function in valid rationalization. These claims may be supplemented by other wider-ranging claims: (vi) physicalism (both token – and type –) is not known to be true and cannot be shown to be true, despite the fact that there is no known argument to show that it must be false; (vii) interactional causal regularities cannot be shown to be subsumable under invariant causal laws; (viii) explanation in science cannot be shown to require subsumption under invariant covering laws, and wherever favored, cannot be shown to be favored for the sake of explanatory truth.

To support any fair part of (i)-(viii) is to subvert Davidson's entire program. The only item in the tally we share with Davidson is (vii), but it is used for opposed purpose here and in Davidson's account. Also, it is entirely possible to proceed to support (i)-(viii) without any developed theory whatsoever of what an action is, and without broaching von Wright's independent objection to theories of Davidson's type, von

6 «Mental Events», p. 208.

Wright's counterclaim strengthens (iv) or replaces it by (iv') rationalization cannot be a form of causal explanation, since (v) is strengthened to, or replaced by, (v') reasons cannot be part of the causes of the actions they explain, since they are not in any sense independent of the actions they may be used to rationalize. (To say this, of course, is not to deny that reasons, the reasons an agent has, may serve as part of the cause of events and effects other than the actions to which they are initially assigned-by reference to which, in addition, actions are first designated as such.)

One begins to see, therefore, what is meant by claiming that the analytic theory of action is utterly impoverished; for it is the case both that the principal reasons for advocating Davidson's sort of theory have very little to do with actually scanning the salient features of what would ordinarily be admitted as exemplars of human action, and that no empirically generous account of the distinctive features of human action has ever been supplied in the analytic philosophy of action – offering a basis for denying any of the claims (i)-(viii). These are very harsh judgments, but the state of the inquiry makes any effort to soften them doubtful at once: unless, that is, such judgments are instantly accompanied by the effective redemption (by argument) of a significant number of denials of the charges collected in (i)-(viii). We ourselves, therefore, must venture beyond our own rhetoric.

II

One counterconsideration to Davidson's theory is this: if (as he says) causality behaves extensionally, then if reasons (what may serve as reasons: beliefs and desires, say) are (as he also says) causes, then they must behave extensionally (which he denies), or they must be able to be treated sufficiently well in an extensional way to be identified as causes at all. If that cannot be done, then the Principle of Causal Interaction must be false. But if, on the other hand, the Principle is true, and if reasons do function sufficiently well extensionally to be identified as causes, then there is no principled reason why they cannot also enter into causal laws (even if, contingently, there are no such valid laws). The «Anomalism of the Mental» would then be false.

Davidson faces an obvious dilemma, therefore, the resolution of which shows the triad to be inconsistent. Alternatively, if the two principles mentioned are, independently, true, the second principle, the «Principle of the Nomological Character of Causality» must be false or dubious (in its turn), because it is inconsistent with the other two or depends, extraneously, on reductionism. That finding would, in fact, be quite promising, in the sense that there is no logically necessary connection between causality and nomologicality; one could point out, for instance, that the very concept of a person entails, on any familiar view, the notion of a capacity to produce intended effects by way of deliberate action within an agent's control; and (contrary to Davidson's view) that there is no logical pressure to hold that such causality must be nomological.

Davidson himself freely declares (in fact, he must do so to support his thesis): «We

are usually far more certain of a singular causal connection than we are of any causal law governing the case»⁷. He is thinking of individual actions. Hence, it would not upset part of Davidson's thesis to admit what may be called an «agential» theory of action – which merely preserves what is entailed by any minimally plausible theory of persons as agents, *plus* what *does* go contrary to Davidson's view, namely, the denial of nomological invariance⁸. That second notion, the notion that causality is nomologically invariant, is, of course, the canonically favored view, the one Davidson prefers – utterly without argument.

We have now established that the triad that forms anomalous monism is irretrievable inconsistent or insuperably doubtful and that an agential conception of causality is both coherent in itself and congruent with ordinary notions of what it is to be a person, an actual human agent, one apt for producing actions. We should also note that there is no reason that Davidson suggests for denying (and he does not deny) that mental states and events do enter into causal interaction; furthermore if actions are either caused by «primary reasons» (in Davidson's sense) or if reasons are logically inseparable from whatever is rightly characterized as an action (in von Wright's sense), then it is impossible to deny that mental states and events enter into causal interaction – however complexly. But if that is true, then it must be false that causality is invariably extensional; and if it is not extensional, then it must be false that causality entails strict nomologicality. Hence, not only is Davidson's triad inconsistent, the Nomological Character of Causality must be a false principle. Furthermore, all this can be determined without raising the question of whether von Wright's implicit denial of any such causal account is correct. The marvel is that all this can be decided without any ramified theory of action at all.

III

A second line of counterargument against Davidson's view concerns his physicalism. Now, there is no known argument to show that a physicalist reduction of mental states and events is incoherent or simply false. There is also none to show that it is true. The best we can say is that physicalist identities, whether type – or token –, are affirmed or denied as a matter of reasonable bet. It needs to be said at once that there are no grounds at all that Davidson provides for betting that a token-identity reading of mental states and events (that is, a physicalism taken token-wise) is true. The only reason he could possibly have for insisting on it is the remote reason of saving anomalous monism as the *only* plausible option *for* saving the unity of science program.

Once that goes (as it must, on the argument supplied), there ceases to be any reason

7 «Reasons, Causes, and Actions», p. 16.

8 See Joseph Margolis, *Science without Unity: Reconciling the Human and Natural Sciences* (Oxford: Basil Blackwell, 1987), ch. 8.

to deny non-extensional causal interaction and non-nomological causality. This is not to say, of course, that causes fail to remain causes under any description; only that there may be no viable way of ensuring that seemingly co-designative descriptions of given actions *are descriptions of one and the same action*. That is just one of the conceptual peculiarities of individuating and identifying actions.

Hence, when, rather famously, Davidson explains:

I flip the switch, turn on the light, and illuminate the room. Unbeknownst to me I also alert a prowler to the fact that I am home. Here I need not have done four things, but only one, of which four descriptions have been given⁹

he leaves us with the severe disadvantage of not being able to specify the criterion on which we (or he) could decide *whether* we are dealing with one action or not. Certainly we could say: take action A; now describe *it* in the four ways Davidson proffers; then, under that constraint, it is true that whatever caused A caused A under any of the four descriptions, and whatever A caused A caused under any of the four descriptions. Causality is extensional to that extent. But, given merely the four descriptions of «an» action in a pertinent context, we lack any rule or principle for deciding whether *those* descriptions identify one action or several different actions (nested perhaps in some noncausal way)¹⁰. There is no way to decide the question, except by offering a theory of what it is to be an action. Davidson never addresses the question.

Davidson's own solution is logically neat but conceptually useless. It commits us to the following thesis:

our primitive actions, the ones we do not by doing something else [that is, by performing some other action], mere movements of the body – these are all the actions there are. We never do more than move our bodies; the rest is up to nature¹¹.

What Davidson means by a «primitive action» is something like pointing one's finger, tying a shoelace, and the like¹². His view is that whatever is an *action* «does require that what the agent does is intentional under some description, and this in turn requires... that what the agent does is known to him under some description»¹³. On this account, actions can be individuated crisply because they can be individuated *non-intentionally*, that is, physically – hence, extensionally.

9 «Reasons, Causes, and Actions», p. 4.

10 See, for instance, Alvin I. Goldman, *A Theory of Human Action* (Englewood Cliffs: Prentice-Hall, 1970). This is not, of course, to endorse Goldman's theory.

11 Donald Davidson, «Agency», *Essays on Actions and Events*, p. 59.

12 He credits Arthur Danto with a similar view. See Arthur C. Danto, «What We Can Do», *Journal of Philosophy*, LX (1963).

13 «Agency», p. 50.

Here, finally, we come to the heart of Davidson's theory of action. For, *if* it can be shown that all actions are individuatable by way of «primitive actions», then the intensional complexity of alternative descriptions can be obviated in *principle*; but *if* that is not possible, then Davidson's program for ensuring the extensionality of causation – both in terms of «primary reasons» and in terms of what may be causally effected by performing an action – will be defeated. Our own claim here is that the «primitive action» thesis fails, though it is certainly not incoherent. It needs to be borne in mind once again that the fate of this particular account does not require resolving the difference between Davidson and von Wright; although it is also true that the plausibility of Davidson's causal theory of «primary reasons» depends very heavily on the plausibility and viability of the «primitive action» thesis – which is incompatible with von Wright's thesis. It is also true that the «primitive action» thesis reduces all intentionally complex description to mere heuristic devices. We should add, before we proceed further, that, in opposing the nomologicality of causality, we are not denying that it may be possible to formulate nomologically exceptionless causal laws. We are insisting only that there is no logical entailment between the two notions and that, in the case of intentionally encumbered actions, there would have to be some range of non-nomological causality. In fact, of course, in recent developments in the philosophy of science, it has become quite problematic to demonstrate that there either *are* invariant causal laws or that the physical sciences require that there be such laws¹⁴, Davidson, ever loyal to Hempel's model, takes no account of these developments¹⁵.

IV

We have prepared the ground for a direct rebuttal of Davidson's theory of «primitive action», that is, the theory that all action may be treated on the model of primitive action. Let us, however, postpone coming to grips with that matter for a moment, and turn instead to tidying up certain loose ends. We have already remarked that mental states and events (if mental events may be admitted) are, normally, intensionally identified. In fact, there is no known principle on which mental states *can* be treated extensionally if either they are not first identified physicalistically or cannot be so identified. For instance, it cannot be shown, on purely extensional grounds, that Tom's thinking that Jack and Jill went up the hill is the same mental phenomenon as Tom's thinking that Jill and Jack went up the hill. Hence if, as seems clear and as Davidson himself admits in his account of «primitive action», actions must be eligible for intentional description in some suitable way if a particular action does really obtain,

14 See, for instance, Nancy Cartwright, *How the Laws of Physics Lie* (Oxford: Clarendon, 1983); Bas C. van Fraassen, *Laws and Symmetry* (Oxford: Clarendon, 1989).

15 See Carl G. Hempel, «Studies in the Logic of Explanation», *Aspects of Scientific Explanation* (New York: Free Press, 1965).

then, whether or not actions are caused by «primary reasons», a fully extensional treatment of action will depend on the fate of the «primitive action» thesis. Of course, the success of the thesis will not save anomalous monism. But that is no longer important. Now, in explaining the intensional behavior of mental ascriptions, in virtue of which psychophysical laws are impossible, Davidson introduces the thesis that rationality is «holistic» – which, of course, affects the conditions under which «primary reasons» are ascribed as the causal antecedents of given actions: *as* that in virtue of which bodily movements are *first* (rightly) open to alterative intentional descriptions. Davidson himself declares quite openly that «My general strategy for trying to show that there are no strict psychophysical laws depends, first, on emphasising the holistic character of the cognitive field»¹⁶:

when we use the concepts of belief, desire, and the rest, we must stand prepared as the evidence accumulates, to adjust our theory in the light of considerations of overall cogency: the constitutive ideal of rationality partly controls each phase in the evolution of what must be an evolving theory¹⁷.

But if, of course, this holism adversely affects «strict» psychophysical laws, *and if* physicalism is false, then the admission of interaction also affects adversely «strict» physical laws. Furthermore, if mental causes taken token-wise are admissible and *if* nomological invariance is indefensible anyway, then no «tight connections»¹⁸ between the mental and the physical are needed either to ensure causality *or* psychophysical laws. So the argument is a red herring.

But if mental states and events are ascribed to agents in accord with the holism just introduced, then, first, it is impossible that token-physicalism be defensible, and, second, it is impossible that the elements of «primary reasons» that provide for explanations by reasons (rationalization) invariably function as causes or function in such a way as to confirm that explanation by reasons is a species of explanation by causes. The reason is quite simple: namely, that, on a holistic model of rationality, it will often be necessary to ascribe particular beliefs or desires and the like to agents in order to yield, or complete, a reasonable sense of the rationality of an agent's behavior, without at all supposing that a particular belief or desire thereby ascribed does obtain determinately in such a way that token-physicalism *could* be invoked. For instance, we often theorize about what beliefs or desires an agent «*must*» have possessed (there may even be disjunctively alternative such attributions to be considered) in order to have acted at all in a fully intentional and rational way. That is, in accord with a holistic view of rationality, explanation by reasons depends on plausibly complete *narratives of rationality*, anchored perhaps here and there (but not everywhere) in relatively determinate

16 Donald Davidson, «Psychology as Philosophy», *Essays on Actions and Events*, p. 231.

17 «Mental Events», p. 223.

18 *Ibid*, p. 222.

belief – and desire – states (that are reflexively and consciously entered into). But if so, then since agents may fairly be said to «have» such and such beliefs and desires even though they cannot function as causes *in Davidson's sense*, it cannot be the case that the reasons an agent «has» (reasonably imputed on grounds of rationality) need also be causes at the same time¹⁹; hence, if reasons are imputably «had» on holistic grounds relative to valid rationalization, despite not being determinate causes, it cannot be the case that explanation by reasons (by reasons «had») is a species of causal explanation (following a Humean or Hempelian pattern). This shows that Davidson's causal theory of action is defective, without at all invoking von Wright's thesis. On the contrary, it provides a very good reason for taking von Wright's view seriously.

V

The rebuttal of the «primitive action» thesis now stands before us in a way relatively freed from extraneous complications. The countermove needed is extraordinarily simple. It is this: even if it is possible to construe a *bodily movement* as a «primitive action» by (first) assigning it (in Davidson's sense) a «primary reason» that functions as a cause of that action (which, on von Wright's theory would at best only correspond to an inseparable intentional feature of that same action), it is not possible to construe all normally admitted actions, first identified in intentionally complex ways, *as primitive actions that we may then describe in intentional terms confined in the general manner Davidson favors*. For, the intentional features of actions cannot always be restricted to the merely *psychological* beliefs and desires and the like internal to the life of the agent. Actions may, to be identified as the familiar actions they are, *have to be ascribed intentional features answering to collective social practices, traditions, institutions that constitute their being just those actions*: that is, *without* its also being true: (a) that those intentional features are, or could be, intra-psychological (merely internal to the psychology of particular agents), or (b) that those intentional features could be construed in terms of type – or token – identity (say, of a neurophysiological sort).

If, for instance, individual agents participate in a revolution or engage in a war or join in a strike, the what, in intentional terms, constitutes what they do *as* those very actions, granting that they have their own psychological reasons and desires? There is no obvious sense in which: (1) the intentional features of their actions could satisfy conditions (a) and (b) (just mentioned), or (2) the actions thus performed could be reasonably identified numerically with any «primitive action». If strategies (1) and (2) fail, then Davidson's (and any similar) physicalism would also fail; *a fortiori*, the «primitive action» thesis (and any similar theory) would fail as well.

The pivotal issue, we may say, is one of conceptual «direction». For, if we identify

19 Cf. «Reasons, Causes, and Actions».

actions as actions, as the particular actions they are, by *first* individuating what agents do intentionally – all the while acknowledging that what they do entails *some* sequence of bodily movements that, however, do *not* enter into the individuation *of* the actions in question (though they do ensure the real occurrence of what is thus individuated) – then it becomes quite problematic whether an action thus identified *is* numerically the same as any putative «primitive action» independently identified by way of suitably congruent intra-psychological beliefs and desires.

What, for instance, is the sense in which *if*, in engaging in a war, a soldier intentionally crooks his finger in the trigger of a gun or picks his nose, the action of being thus engaged is numerically identical with the action of crooking his finger or picking his nose, or identical with any extended sequence of such mere «primitive actions»?

There is a directional asymmetry between construing a primitive action as identical with *that same action* whenever attributed more and more complex intentional descriptions and construing a complex action *first* identified in contextually pertinent intentional terms (without conceding that it is individuated by reference to the enabling primitive bodily movement or movements it is manifested in) as identical with *that or any primitive action*. Davidson addressed only the first possibility; he never discusses the second.

Obviously, we must confine ourselves to the general strategies of argument. (We cannot do justice here to all the issues those strategies provoke.) *If* then, the intentional descriptions that mark individuated actions as the actions they are, are *not* able to be restricted to the intra-psychological feature of (what Davidson calls) the «primary reason» for an action, *then not all actions can be numerically identified with primitive actions*, even if primitive actions be admitted or form ingredients of such actions. In particular, if the intentional description by which some actions are individuated include *collective features* – that is, features that may be assigned to the actions of individual agents but that cannot be construed (or reduced or interpreted solely *in intrapsychological terms*) – then the «primitive action» thesis must be false.

The point at stake is indifferently Hegelian, Marxist, Wittgensteinian, hermeneutic, Frankfurt Critical, structuralist, Foucauldian. That is, the intentional import of what people do is invariably qualified by the collective social practices they share – by the structured forms of life they share – in accord with a division of social labor, in virtue of which no one can have (or need have) internalized *psychologically* the entire set of such practices that give their psychological states the intentional force they have; also, on occasion, what people *do* may make no sense at all except *as* the individuated parts of collective actions (even though, of course, there are no collective agents – only aggregated individual agents, who may even believe they serve or form part of a collective agent). Thus, for instance, *offending another's honor* may well depend, intentionally, on features of a collective form of life that a particular agent is unaware of; and *engaging in a revolution* invariably intails features that cannot be analyzed solely in intra-psychological terms even though they may be assigned to the behavior of suitably endowed individual agents. For example, the notorious case of Zande witch-

craft is utterly unintelligible without reference to the collective life of the Azande²⁰; furthermore, individual actions that instantiate the patterns of bourgeois economy are, irreducibly, elements of a collective practice, as Marx forever reminds us²¹.

Davidson never shows us how, or why, we should treat actions in accord with his «primitive action» thesis. He cannot possibly hold to his view unless he utterly discounts historically shifting divergent forms of life: unless, that is, he insists that what alone makes action initially intelligible as action is whatever, intra-psychologically, rationalizes some «primitive action» somehow identified with the socially complex actions we first identify. But there seems to be no motivation for this except the demonstrably failed objective conforming with the unity of science program; and there is no demonstration offered for plausibly reducing all familiar discourse about action to any ordered account of «primitive actions». Are there, for instance, fatal distortions in understanding action that would result? Davidson never considers the question.

VI

There are, in fact, at least two ubiquitous features of actions not obviously reconcilable with Davidson's thesis, that are captured by the following considerations: (A) we usually describe, explain, individuate, even reidentify actions without regard to «primitive actions» and even without close regard to particular bodily movements (of whatever complexity), so long as *some* suitable bodily movements do obtain and are sufficient for ensuring that the actions in question do occur; we are even noticeably tolerant of descriptive mistakes about such movements, or indifferent to their particular details, and much less tolerant about mistakes regarding intentional features of actions that include such movements; and (B) there are no systematic or reliable inferences that go in either direction, token – or type – wise, between mere bodily movements involved in actions (whether suitable for «primitive actions» or not) and intentionally complex actions manifested in or by them. (A) and (B), neither jointly nor separately, show that Davidson's theory is incoherent or impossible to defend; but, together, they do show that it would be an unreasonable bet to support his thesis without first offsetting their adverse significance. In fact, (A) and (B) point to a curiously deep puzzle that Davidson never addresses, namely, how in general, do we recognize and identify an action as opposed to a mere bodily movement. The question, of course, is Wittgenstein's.

Consider this: even if it were the case that we first identify a bodily movement suitable for a «primitive action» and then impute, where pertinent, complex intentional

20 See Peter Winch's much-discussed paper, «Understanding a Primitive Society, *American Philosophical Quarterly*, I (1964).

21 See, for instance, Karl Marx, *Grundrisse*, trans. Martin Nicolaus (New York: Vintage, 1973), pp. 105-106. Alex Callinicos has attempted, not quite satisfactorily, to reconcile Davidson's with Marx's analysis of action. See his *Making History* (Ithaca: Cornell University Press, 1988), ch. 3.

properties to the action thus individuated, we should have to face the question: Which criteria determine when and on what grounds such *redescriptions* remain «objective» where they *depart* from the «primary reason» with which a given agent claims to have performed the action in question? There must be some pertinently contextual grounds on which intentional descriptions of apparent actions are thought to be reasonably objective. For example, a young boy acting in accord with a socially well-defined *rite de passage* – becoming a man, say – could be described as having performed this or that particular (intentionally complex) act, or as having produced this or that action, whether or not he was aware of what he was doing, in the sense of being able to supply, conformably, a «primary reason» of his own; such a description might well take precedence over anything *he* might volunteer as his own primary reason; his grasp of what he had done might also be objectively assessed in terms of his understanding *such* a societally entrenched description; and, in any case, *that* he had performed this or that particular act or action might depend on the weight of such societally entrenched descriptions rather than merely on its being true that his own would-be «primary reason» *caused* the (initial) «primitive action» alleged.

In other words we normally do not identify an action as the action it is by reference first, or only, to a causal «primary reason» that produces a «primary action», even when it may be possible to construe things thus – for at least two reasons: first, because it is normally not necessary to identify an action as an action, *by* identifying it as having been caused, or as having been caused by a «primary reason» (even when the agent is prepared to volunteer information regarding what he believed and desired *in* performing the action he performed); and second, because it is entirely usual to identify an action by reference to societally entrenched intentional descriptions even if the agent would not favor those descriptions or claim them as his primary reason (provided only he had some suitable awareness of what he was doing). These caveats have interesting implications.

First of all, Davidson must disallow (implausibly) that an intentional description of an action may «objectively» *override* or take precedence over an agent's «primary reason». Hence, secondly, he must (implausibly) disallow that an action may, in principle, be identified and individuated in intentional terms, «objectively» *without* reference at all to the causal force of a would-be «primary reason». Thirdly, he must disallow that actions individuated in intentional terms may yet be numerically *different* from one another even if they involve the same primitive action or the same bodily movement (that could in principle be identified as a primitive action). Fourthly, he must (without argument) either disallow that «non-primitive actions» (those that, on his own account, «can be explained as caused by another event [action] of which he is [the] agent»²² are *numerically different* from *any* (other) primitive action, or he must lack altogether an «objective» basis for determining the proper intentional description of *such* «non-primitive actions». Fifthly, he must (again, without argument) disallow that «non-

22 «Agency», p. 49.

primitive actions» as well as primitive actions may be «objectively» individuated *without* reference at all to the causal conditions that produce them. And sixthly, he must (again implausibly) refuse to make a relevant distinction between «primitive actions» and numerically distinct «non-primitive actions» performed by performing one of his primitive actions – but not produced *by* any causal linkage initiated *by* such primitive actions²³.

VII

Now, all of these difficulties may be obviated at a stroke by abandoning the causal theory of action Davidson espouses. For, on Davidson's view, as a Humean or Hempelian, causation entails the logical independence of cause and effect; in context²⁴, causation entails the logical independence of certain mental states or events and behavior that rightly counts as action. It is entirely possible, however, to distinguish, logically, between mental states and behavior, without holding that the mental states (or events or onset of states) that (on Davidson's view) *cause* any behavior that rightly counts as action *are ever actually independent of the particular actions they are thought to produce*.

At this point, the argument against Davidson is surprisingly straightforward. For, in proposing his thesis about «primary reasons», Davidson begins by claiming only that «it is possible» to construe rationalizations as causal explanations. His argument to this end involves two theses:

1. In order to understand how a reason of any kind rationalizes an action it is necessary and sufficient that we see, at least in essential outline, how to construct a primary reason.
2. The primary reason for an action is its cause²⁵.

Davidson does not actually say here – though he does believe – that it is «necessary and sufficient», if we are to rationalize an action, that we be able to construct a «primary reason» for it, and that it is *necessary and sufficient* that «the primary reason for an action *be* its cause».

There is nothing, however, that Davidson says anywhere that *shows* that the second claim is true. Its denial is not self-contradictory – is, on the contrary, theoretically advantageous, as we have just seen. Davidson does say that «to know a primary reason

23 Cf. *Ibid.*, p. 59.

24 See Frederick Stoutland, «Basic Actions and Causality», *Journal of Philosophy*, LXV (1968); and «The Logical Connection Argument», in Nicholas Rescher (ed.), *Studies in the Theory of Knowledge* (Oxford: 1970).

25 «Reasons, Causes, and Actions», p. 4.

why someone acted as he did is to know an intention with which the action was done»²⁶. But that formulating is entirely neutral as between the causal and non-causal reading (where, by the latter, we mean that a reason for acting is logically inseparable from the *action* performed – not merely from admissible *descriptions* of that action). Davidson *does* (very briefly) *appear* to claim that thesis 2 (cited above) is necessarily true independent of his theory linking rationalization and causal explanation. For he says: «In order to turn the first ‘and’ to ‘because’ in ‘He exercised *and* he wanted to reduce and thought exercise would do it’, we must, as the basic move, augment condition C1 [thesis 1 of the citation above] with:

C2. A primary reason for an action is its cause²⁷. That is, we *must* not merely to square the account with the thesis that rationalization is a species of causal explanation but to square it with the very nature of a «primary reason». The denial of this second thesis is, of course, very nearly the main burden of von Wright’s account. For von Wright sides with the «intentionalists» against the «causalists», in holding that «the connection between intention and behavior [is] of a conceptual or logical nature».

von Wright is explicit about his «not denying that, e.g., desires or wants could have a causal influence on behavior. Nor am I disputing [he adds] the obvious role which dispositions, habits, inclinations, and other regularities and uniformities of behavior play in the explanation and understanding of action»²⁸. But he makes it clear as well that what, in effect, is the substance of Davidson’s «primary reasons» is abstracted from what (on his own view) is logically and substantively internal to the onset of the very action those reasons are supposed (on the «causalist’s» view) to cause; and that the causal model is committed (as the «intentionalist’s» model is not) to explanatory subsumption «in accordance with the covering-law model»²⁹.

So the «intentionalist» alternative enjoys two extremely important advantages: for one, it solves at once the puzzle of how we individuate and identify the actions we do – in a way that conforms with familiar practice; and for another, it is not burdened with the assumption that the relevant intentional factors of belief and desire either enter as such – or can be matched tokenwise with physical events that enter as such – into lawlike connections apt for explaining the actions in question. It is, of course, notoriously difficult, probably impossible, to show that the functional or «folk» distinctions of reasons and desires *could* ever be suitably linked with nomological invariances³⁰.

Our argument is now complete, at least as far as demonstrating the following: first, that Davidson’s theory rests on, and includes, untenable claims; second, that where it is at least self-consistent, it is arbitrary and produces avoidable conceptual puzzles that it

26 *Ibid.*, p. 7.

27 *Ibid.*, pp. 11-12. The claim is completely in accord, on this reading, with «Freedom to Act», *Essays on Actions and Events*.

28 *Explanation and Understanding*, pp. 95-96.

29 *Ibid.*, pp. 97-98.

30 See, for instance, Jerry A. Fodor, *The Language of Thought* (New York: Thomas Y. Crowell, 1975), Introduction.

never addresses; third, that it is not in any demonstrable respect superior to the («intentionalist») theory that it was intended but failed to defeat; and fourth, that it does not address the most salient features of our discourse about action.

The irony is that it has taken more than a quarter of a century to undo the mischief of Davidson's model. It will continue being influential, of course, but the rest of us may in good conscience move on to a very different basket of questions that a genuinely ramified account of action should address – that «analytic theories» should now address – since those questions have been explored for a long time (in an entirely different idiom) by the various Continental European traditions.

A second glance suggests that the defeat of Davidson's entire line of analysis is less a matter of small adjustments to a research program that had correctly oriented the Anglo-American philosophical tradition than the wholesale rejection of a comprehensive line of thinking too closely allied with the unity of science program.

Now, the philosophy of action is converging upon the profoundly historicist nature of the social world. The very issues Davidson ignored – (a) that of the conditions under which actions are rightly characterized by this or that intentional description, (b) that of the constraints under which actions viewed in terms of one such description may be justifiably reinterpreted and ascribed a changing or evolving significance, (c) that of the relationship between causal and narrative forms of understanding, (d) that of the bearing of collective social practices on the formation of agency, (e) that of how science arises as a form of social *praxis*, (f) that of the historically constructed and variable nature of human agents, (g) that of the analysis of human existence, (h) that of the metaphysics of self-understanding – all these questions, and more, suddenly loom as familiar companions, that an earlier enchantment had kept us from perceiving. Our farewells, then, look away from our own impoverished past.

Joseph Margolis
Department of Philosophy
Temple University
College of Arts and Sciences
Humanities Building 022-32
PHILADELPHIA
PENNSYLVANIA 19122
(ESTADOS UNIDOS)