

INVITED PAPER

Hans-Johann Glock

Reasons for Action: Wittgensteinian and Davidsonian Perspectives in Historical and Meta-Philosophical Context

Abstract

My paper reflects on the debate about reasons for action and action explanations between Wittgensteinian teleological approaches and causalist theories inspired by Davidson. After a brief discussion of similarities and differences in the philosophy of language, I sketch the prehistory and history of the controversy. I show that the conflict between Wittgenstein and Davidson revolves neither around revisionism nor around naturalism. Even in the philosophy of mind and action, Davidson is not as remote from Wittgenstein and his followers as is commonly assumed: there are numerous points of contact of both a biographical and a substantive kind. The real conflict concerns the difference between Davidson's official subjectivist approach to reasons for action, according to which they are mental states of believing and desiring (or their onsets) on the one hand, and an objectivist approach, according to which reasons for action are what is believed or desired, non-mental facts or states of affairs.

1. Wittgenstein and Davidson: Similarities

2013 marked the 60th anniversary of the posthumous publication of Wittgenstein's *Philosophical Investigations*. It also marked the 10th anniversary of the death of Donald Davidson and the 50th

anniversary of his seminal article “Actions, Reasons and Causes” (ARC): a fitting occasion to reflect on the relation between Wittgenstein and Davidson. Several areas invite comparison (Glock 2003: 1-39 and 2013). There are interesting parallels at the strategic level. In different ways, both are part of the linguistic turn of analytic philosophy: they ascribe a central philosophical role to language, albeit for different reasons. Furthermore, they propound conceptions of language and the mind that shun both Platonism and Cartesianism. Linguistic expressions acquire meaning not by being associated with either abstract entities or private mental processes, but by having a certain role or function. More generally, in both Davidson and the later Wittgenstein there are important themes that are loosely speaking pragmatist, notably the stress on the philosophical importance of human action, a holistic perspective that views individual expressions, sentences and thoughts as part of a larger context of activity, and a communitarian tendency to view that context as inherently social. Finally, both share an anthropological approach to philosophy. Even their discussions of topics apparently unrelated to human affairs – e.g. truth, modality, mathematics, causation – revolve ultimately around a *philosophical anthropology*, a conception of human behaviour and human capacities in general, and of language and linguistic capacities in particular.

In the later Wittgenstein, there is a pervasive, albeit often implicit, anthropological theme. It comes to the fore in the central role he assigned to “forms of life” (see Glock 1996: 124-129). Wittgenstein believed that an “ethnological point of view” envisaging different types of human practices helps philosophers to see “matters more objectively” (CV 37).

He used to say that what we might call “the anthropological method” had proved particularly fruitful in philosophy: that is, imagining “a tribe among whom it [a certain human activity] is carried on in this way: ...” (Rhees 1965: 25)

The central place of human beings and human action in Davidson’s work has been noticed less. Yet his heuristic device of “radical interpretation”, though derived from Quine’s “radical translation”, is connected to Wittgenstein’s anthropological method. Radical

interpretation is the attempt to understand an alien community without the benefit of any prior knowledge of what its members think or what their utterances mean. Following Quine's indeterminacy of translation, Davidson argues that there are equally correct though mutually incompatible ways of making sense of the interpretees. Whereas Wittgenstein envisages exotic practices, e.g. measuring lengths through elastic rulers, Davidson envisages deviant interpretations, e.g. rendering alien sentences as statements not about material objects but about their centres of gravity. Both procedures, however, raise the spectre of *alternative* ways of thinking and acting that have hitherto gone unnoticed.

To be sure, this is not the purpose for which Davidson employs radical interpretation. By considering the anthropological scenario, he instead hopes to shed light on the master-problem of his thought: What is it to understand other human beings? His solution is provided by a unified theory of meaning and thought which intertwines the philosophy of language, the philosophy of mind and the theory of action (2004: ch. 10). Understanding what a person does, what she believes and desires, and what she means by her utterances are part of a single enterprise. The aim is to understand human behaviour by explaining it causally. Nevertheless, the picture of human beings that emerges has strong Kantian affinities. Although we are part of the causal order, we must also treat each other as rational agents. For, following Davidson's well-known "principle of charity", the only way of making sense of the mental states of agents, their linguistic utterances and their actions is to regard them as being by-and-large rational.

There are also similarities between Wittgenstein and Davidson at a tactical level. One is a striking affinity between the *Tractatus* and Davidson's idea of a systematic theory of meaning for natural languages. The immediate roots of this project lie in the logical semantics of Tarski. However, the languages Davidson is interested in are natural rather than artificial. In this respect, he stands in the tradition not of Tarski and Quine, but of Wittgenstein's *Tractatus*. Formal logic does not furnish an ideal language which avoids the shortcomings of ordinary language, it serves to indicate the

underlying logical form that sentences in the vernacular possessed *all along*. Contrary to received wisdom, therefore, neither the early Wittgenstein nor Davidson are ideal language philosophers.

Quine's naturalism lays claim to constituting the revisionary metaphysics of modern science. By contrast to his teacher and friend, however, Davidson aims to bring out the "metaphysics implicit in natural language". He is interested not in "improving on natural language, but in understanding it". Alluding to a simile of the later Wittgenstein, he describes "the language of science not as a substitute for our present language, but as a suburb of it" (1984: 203; 1985b: 172, 176). Again in line with the later Wittgenstein, Davidson subscribes to a third-person perspective on language. More specifically, he links the linguistic meaning of an expression to its use.

"[M]eaning is use", quoth Wittgenstein. The idea is obvious, but its full force is mostly unappreciated or misappropriated. Misappropriated by those who would convert any typical purpose served by uttering a sentence into a kind of meaning. Unappreciated by those who treat Wittgenstein's slogan as gesturing at a way of discovering a meaning already embedded in an expression. What wants emphasizing is not that use points the way to preexisting meanings, but that it creates, and so constitutes, meaning. (1999b: 80)

From a bird's eye view of contemporary analytic philosophy, Wittgenstein and Davidson may appear comparatively close, at least relative to the revival of essentialist metaphysics on the one hand, the naturalistic dissolution of philosophy into research programmes in cognitive science and AI on the other.

2. Wittgenstein and Davidson: Differences

The differences are equally important. The later Wittgenstein became highly sceptical about the philosophical value of logical analysis, notwithstanding the aforementioned fact that his own *Tractatus* was one of the major inspirations behind formal semantics. By contrast, following Tarski and Quine, Davidson sets great store by the power of formal logic. But then again, like Quine, Davidson combines logical analysis and formal semantics à la

Tarski and Carnap with a pragmatist emphasis on language as a form of human behaviour.

Next, whereas Wittgenstein's conception of the connection between meaning and use is *normativist* and *communitarian*, Davidson's is *anti-normativist* and *individualist* (Glock 2010). Both agree that expressions do not have an intrinsic linguistic meaning. Rather, what they mean is determined by how they are used.

... let's not forget that a word hasn't got a meaning given to it, as it were, by a power independent of us, so that there could be a kind of scientific investigation into what the word *really* means. A word has the meaning someone has given to it. (BBB 28)

... when we learn or discover what words mean, the process of learning is bestowing on words whatever meaning they have for the learner. It's not as though words have some wonderful thing called a meaning to which those words have somehow become attached, and the learning process is just putting us in touch with that meaning. (Davidson 1999b: 41)

However, Davidson turns the dependence of meaning on use into an argument against the Wittgensteinian idea that meaning is constituted by shared rules of use: if the meaning of an expression *e* in a language L is determined or constituted by the use speakers make of *e*, the meaning of *e* cannot at the same time prescribe a certain use to the speakers of L. Accordingly, the idea that the use of *e* is governed by lexical norms is incompatible with the idea that meaning is determined by use. It is a lapse into the Platonist picture according to which "the meanings of words are magically independent of the speaker's intentions" (Davidson 1990a: 310). Wittgenstein and his followers would answer this complaint as follows: We must distinguish between the use an individual speaker *a* makes of *e*, and the use that the linguistic community *C* makes of it. *Communal* use *constitutes* meaning, while *individual* use is *responsible* to it. By this token, the existence of lexical norms and their independence of the utterances and intentions of individual speakers is no more mysterious than the existence of legal norms and their independence of the acts and motives of individual agents.

For related reasons, the social dimension of language that our protagonists diagnose is strikingly different. Wittgenstein's private language argument purports to demonstrate the incoherence of the idea of a language that is not subject to semantic rules that *can be shared* between different speakers, without excluding the possibility of a *solitary language* that a speaker uses for purposes other than communication. Davidson's triangulation argument aims to show that thought and meaning require *actual communication*, but without requiring that this communicative employment be guided by rules, let alone rules shared between speaker and hearer. As a result, he regards language as a *sum of idiolects*, whereas for Wittgenstein it is a *normative practice* shared between the members of a linguistic community.

Wittgenstein and Davidson are closer in the results they reach through, respectively, the anthropological method and radical interpretation. Both insist that there are *minimal requirements* which a form of linguistic behaviour must meet in order to be intelligible to us. According to the principle of charity, interpretation presupposes that we can treat the aliens' beliefs as by and large true (Glock 2003: 194-199). Wittgenstein concurs partly. "If language is to be a means of communication there must be agreement not only in definitions but also ... in judgements" (PI §242). But while the principle of charity stresses the second point, it discards the first. By maximizing agreement in opinion it puts the cart – truth – before the horse – meaning. By and large, we must understand *what people say* in order to judge whether they are speaking the truth. The consensus presupposed in sharing a language is ultimately "not agreement in opinions but in form of life" (PI §241; see RFM 353). By the same token, understanding an alien language relies on convergence not of *beliefs*, but of patterns of *behaviour*, which in turn presupposes common perceptual capacities, needs and emotions: "The common behaviour of mankind is the system of reference by means of which we interpret an unknown language" (PI §206).

Finally, both Wittgenstein and Davidson have interesting yet disparate views on the topic of animal minds and animal thought. Both approach mental phenomena from a third-person perspective. They do not appeal to phenomena – whether mental

or neurophysiological – that cannot be manifested in behaviour even in principle. Both hold in effect that one cannot attribute beliefs to creatures who are totally incapable of manifesting these beliefs, Davidson because he cannot make sense of the notion of a belief as a private attitude completely detached from behaviour and its explanation (1984: 170; 2001: 99-100), Wittgenstein because he insists that we can ascribe a thought that p to a creature a only if something counts as a thinking that p rather than that q . But there is also a crucial difference. Unlike Wittgenstein, Davidson insists that even for simple beliefs, the required behaviour must include *linguistic* behaviour. Consequently he denies that non-linguistic animals have any thoughts. Wittgenstein (by contrast to some of his disciples) adopts a position closer to common sense – for once, many might be tempted to say. Animals are capable of having beliefs, desires and intentions of a *simple* kind, namely those that can be expressed in non-linguistic behaviour (Glock 2000).

3. Actions, Reasons and Explanations

This “compare and contrast” exercise could be continued on several other topics. My paper is devoted to one of them. It is the nature of action, and, more specifically, of reasons for action, and of action explanations. This cluster of issues is particularly topical because of the aforementioned anniversary of Davidson’s most famous article. It is also essential to interpreting and comparing our two protagonists, given the central place they accord to human behaviour. Finally, agency is an especially fruitful subject from a philosophical point of view. The debate between, loosely speaking, Wittgensteinian teleological approaches and causalist theories inspired by Davidson has been reopened over the last 20 years. In the process, novel light has been thrown on age-old problems that mainstream theorists of action had regarded as settled.

The three central problems in the theory of action are:

1. What are (intentional, voluntary, rational, etc.) actions? (The nature of actions.)
2. What are the reasons for which (intentional, voluntary, rational, etc.) actions are performed? (The nature of reasons for action.)

3. How are actions to be adequately explained? In particular, what is the nature and status of explanations of actions by reference to the reasons for which agents perform them? (The nature of action explanation.)

The problem of the nature of actions revolves around a set of interconnected differences:

- the difference between what agents do and what merely happens;
- the more specific difference between what agents do and what happens to them (what they undergo);
- the difference between what agents do and what they deliberately refrain from doing (acts vs. omissions);
- the difference between what agents do, and what they merely cause or bring about, e.g. what they intend to do and its unintended side-effects;
- the difference between intentional actions and mere behaviour, e.g. automatic or unconscious reactions or the behaviour of (all or many) non-human animals.

As regards the first contrast, actions are standardly conceived as a *species of events*. There are alternative proposals which assign actions to other categories, or regard them as *sui generis*, constituting a category of their own. These alternatives are incompatible with a major contention of Davidson's; and some of them (notably von Wright 1963: 35-7; 1971: ch. II; see Stoutland 2010) – were inspired by Wittgenstein. Nevertheless, for the sake of exposition and argument I shall assume that actions are events.

This establishes a minimal common ground with the causal theories epitomized by Davidson. They treat actions as *events with a special kind of cause*. Causal theories of action – causalism in short – distinguish themselves from other positions – notably those invoking agent causation – through their stance on the other two problems – the nature of reasons for actions and the nature of explanations of actions by reference to such reasons--what Davidson calls “rationalizations” and I “intentional explanations”. According to causalism about reasons, the reason for which an agent A acts is an *inner mental cause of A's outer bodily movements /*

behaviour. According to causalism about explanations, the intentional explanation of A's action is an explanation in terms of *efficient causes*, just like explanations of physical events.

It is sometimes suggested that causalism about reasons is a “stronger” and causalism about explanation a “weaker claim” (Alvarez 2007: 105). Accordingly, causalism about reasons *implies* causalism about explanations, yet *not vice versa*. Both claims are to be taken with a pinch of salt. As regards the second, one can deny that reasons themselves are causes of actions, while insisting that intentional explanations make reference to efficient causes *associated with* reasons (see sct. 14). As regards the first: if reasons are efficient interior causes of actions, explaining the latter by reference to reasons is indeed a species of causal explanation, in the sense that it *talks about* efficient causes. Nonetheless *the way we talk about* reasons and actions in intentional explanations might *differ* from the way we talk about the relation between other efficient causes and their effects. The terms referring to actions and reasons in intentional explanations and the terms referring to events in other causal explanations might have the same reference yet differ in Fregean sense. Causalism about explanation does require, however, that the relation we attribute is the same in all cases, namely one of efficient causation.

4. The History of the Debate

Causalism has been the default position in modern philosophy. From Descartes through Hobbes, Locke, Berkeley, Hume and Mill to Prichard, actions were regarded as bodily movements caused by interior acts of the will (“volitions”). Volitionism was decisively refuted by Wittgenstein and Ryle (Candlish & Damnjanovic 2010).¹ This did not spell the end for causalism, however. As regards reasons for action, mentalist versions referring to acts of a mind or soul were replaced by materialist versions, according to which the

¹ One related question remains controversial, namely whether all actions involve an element of *trying*. When we Φ , do we always *try* to Φ ? This is claimed by Hornsby (1980) and Grice (1989). For a Wittgensteinian rebuttal of the Gricean defence of this position by appeal to conversational implicatures see Glock 1996.

reasons/causes of action are neurophysiological phenomena. As regards the explanation of action, there had been a prolonged *Erklären-Verstehen* (explaining-understanding) controversy in the 19th century. It pitted positivists and empiricists like Mill against the hermeneutic tradition and neo-Kantianism. The hermeneutic side led by Dilthey insisted that the explanations of actions in the social and historical sciences is a *sui generis* form of understanding which is fundamentally distinct from the causal explanations in terms of efficient causes furnished by the natural sciences.² Logical positivists, led by Hempel (1942), combated this methodological pluralism in the name of the unity of science. They denied that there is a fundamental methodological difference between the explanations of actions by reference to reasons and the causal (deductive-nomological) explanations of natural science.

Ryle's *The Concept of Mind* and Wittgenstein's *Philosophical Investigations* marked a turn against logical positivism and causalism. They fuelled the rise of an analytic variant of hermeneutics (von Wright 1971: 29-30, 181-2). Analytic hermeneuticians like Dray, Winch, von Wright and Taylor tried to reinstate the dichotomy between causal explanation and the understanding of actions. Wittgenstein also inspired thinkers like Anscombe, Kenny and Melden. They were interested less in the methodology of the humanities and social sciences and more in moral psychology, especially the problem of freedom, the former two from an Aristotelian-cum-Thomist perspective.³ Both the analytic

² A different though related methodological distinction was drawn by neo-Kantians like Windelband and Rickert, namely between the natural sciences which are “nomothetic”, i.e. concerned with general laws and the social/historical sciences which are “ideographic”, i.e. concerned with individual cases. This dichotomy is connected with the issue of psychological laws.

³ Aristotle himself has been claimed for the causalist camp (e.g. Davidson 1980: 11). After all, he contends that “the origin of action – its efficient, not its final cause – is choice, and that of choice is desire and reasoning with a view to an end” (*Nicomachean Ethics* 1139a 32ff.). The attribution has been contested by Alvarez 2007: 113-4n. She invokes *Nicomachean Ethics* 1139 b2-3. “Hence choice is either desiderative reason or ratiocative desire, and such an origin of action *is a man*”. But this at best shows that Aristotle thought that the efficient causes of action include both a choice by the agent and the agent herself – a version of agent causation. Such a position is of dubious coherence: i) it portrays the action as causally overdetermined; ii) it assimilates two rather distinct relations in which

hermeneuticians and the moral psychologists propounded teleological accounts and opposed causalism: the reasons for which agents act are not interior events causing the action, and intentional explanations are *toto caelo* different from causal explanations.

During the 1950s and the early 1960s, teleological approaches dominated the theory of action, especially in Britain. This short-lived Wittgensteinian orthodoxy was epitomized by (though far from confined to) a series entitled *Studies in Philosophical Psychology* (see Sandis 2015). Looking back, Davidson wrote.

In December of 1961 Hempel gave the presidential address at the annual meeting of the Eastern Division of the American Philosophical Association. The title was “Rational Action”. In that address, Hempel argued that explanation of intentional action by appeal to the agent’s reasons does not differ in its general logical character from explanation generally; in taking this position, he was swimming against a very strong neo-Wittgensteinian current of small red books. (Davidson 1980: 261)

Just in case you thought that Davidson might have had Mao bibles in mind, he stated the aim of “Actions, Reasons and Causes” as follows:

In this paper I want to defend the ancient – and commonsense – position that rationalization is a species of causal explanation. The defence no doubt requires some redeployment, but it does not seem necessary to abandon the position, as has been urged by many recent writers [n: Some examples: Gilbert Ryle, *The Concept of Mind*, G.E.M. Anscombe, *Intention*, Stuart Hampshire, *Thought and Action*, H.L.A. Hart and A.M. Honoré, *Causation in the Law*, William Dray, *Laws and Explanation in History*, and most of the books in the series edited by R.F. Holland, *Studies in Philosophical Psychology*, including Anthony Kenny, *Action, Emotion and Will*, and A.I. Melden, *Free Action*. (Davidson 1980: 3)

There is no doubt that Davidson succeeded in stemming the “neo-Wittgensteinian” tide of “small red books”. His article reinstated causalism not merely as the majority view, but as something that is widely regarded as a truism. The version of causalism inspired by

the action stands, namely to a choice and to the agent, lumping both together under the label of efficient causes.

Davidson is known as “the standard story” (Smith 2004). According to this story, actions are bodily movements that are both explained and rationalized by inner – mental or neurophysiological – states or events. ARC owed its success to a powerful combination of rebuttals of then popular anti-causalist arguments and novel considerations in favour of causalism. To its mainstream proponents, the fact that causalism rose like Phoenix from the ashes simply reflects the compelling nature of the arguments – both negative and positive – propounded by Davidson and his followers. Unsurprisingly, opponents of causalism reject that diagnosis. Although few of them recognize this, they owe us, among other things, an alternative explanation of why causalism caught on as a renewed orthodoxy. The next section discusses an interesting attempt to meet that obligation.

5. Teleology vs. Causalism – a Battle in a “Meta-Philosophical War”?

Among contemporary Wittgensteinians, it is a commonplace that the influence of Wittgenstein on analytic philosophy has declined sharply since the 1960s, and that many of his lessons have been forgotten. It is equally common to blame this melancholy development not on the quality of specific objections raised against Wittgenstein’s ideas, but on misunderstandings on the one hand, general philosophical trends like the rise of naturalism on the other (Glock 1996: 29 and, more elaborately, Hacker 1996: ch. 8). Regarding the specific case of the waxing of Davidsonian causalism and the waning of Wittgensteinian teleology, D’Oro has recently furnished a similar diagnosis.

In the following I argue that the success of causalism cannot be fully accounted for by considering the outcome of first-order debates in the philosophy of action and that it is to be explained instead by a shift in meta-philosophical assumptions. It is the commitment to a certain second-order view of the role and character of philosophical analysis, rather than the conclusive nature of the arguments for causalism, that is largely responsible for the rise of the recent causalist consensus. (2012: 207)

According to D’Oro, the “philosophical battle” between causalists and teleologists is part of a “meta-philosophical war”, and its outcome depends on that of this wider conflict. More specifically, it depends on the outcome of a conflict between what Strawson famously labelled, respectively, descriptive and revisionary metaphysics. Teleological accounts, D’Oro contends, pursue a descriptive project, while a revisionist orientation is taken for granted by causalists. Accordingly, the “official story” of the triumph of causalism is incorrect. Causalism cannot be supported by “Davidson’s master-argument”, as the best explanation of the distinction between *a* reason A had to Φ and *the* reason for which A *actually* Φ -ed. It could only be established if its revisionist meta-philosophy could be vindicated against the descriptivist alternative. And D’Oro leaves little doubt that she regards this as impossible, since descriptive metaphysics is in the right and revisionary metaphysics in the wrong (2012: 220-1).

The idea of a profound meta-philosophical rift appears *prima facie* plausible.⁴ D’Oro points out that Collingwood was both a descriptive metaphysician *avant la lettre* and an inspiration for analytic hermeneuticians like Dray (2012: 217-20). A meta-philosophical explanation of the philosophical controversy also chimes with the contrast between the Wittgensteinian roots of teleology on the one hand, and the Quinean roots of causalism on the other. Davidson’s rejection of teleology seems to fit in with other naturalistic attacks on a tenet uniting Wittgenstein with conceptual analysis, namely that there is a fundamental divide between science and philosophy, factual/empirical and conceptual/a priori investigations, notably Quine’s repudiation of the analytic/synthetic distinction and Putnam’s attack on the criteria/symptom dichotomy. Finally, D’Oro rightly points out that causalists like Jaegwon Kim take issue with certain methodological-cum-meta-philosophical ideas and attitudes of teleologists. Nevertheless her meta-philosophical diagnosis is inaccurate in several respects.

⁴ This paragraph is indebted to Tripodi (2014). But my meta-diagnosis of the meta-philosophical diagnosis differs from his and is more sceptical.

For one thing, causalists like Kim do not reject teleological approaches simply on meta-philosophical grounds, independently of “first-order” philosophical arguments like that of Davidson. Their *meta*-philosophical complaints are a mirror image of D’Oro’s own: they are designed to explain and criticize the refusal of the opposition to see the *philosophical* light.

For another, the descriptivism D’Oro attributes to the teleologists and subscribes to herself is an uneasy mixture of at least three distinct ideas (2012: 210-212, 217). First, there is the descriptivist project of delineating our actual conceptual scheme instead of revising it according to the presumed dictates of the nature of reality. Secondly, there are “common-sense” claims, presumably considered beliefs firmly held by laypeople, as opposed to the views advanced by scientists, philosophers, religious leaders, etc. Thirdly, there is “folk-psychology”, our everyday mental discourse and the mental concepts it employs, notably that of beliefs, desires and intentions, as opposed to the neurophysiological discourse that eliminativist materialists would like to put in its place. Contrary to a commonly held view, folk psychology does not constitute a scientific or proto-scientific theory of human behaviour. Unlike common sense, it is a network of concepts rather than beliefs (Glock 2013). What is more, common-sense views need not conform to the findings of descriptive metaphysics / conceptual analysis, since they may be guided by an inadequate grasp of the conceptual landscape – by the “surface” rather than the “depth-grammar” of expressions (PI §464). It may well be commonplace, for instance, to regard the mind as a thing and having a pain as a relation of ownership to a private sensation. For better or worse, conceptual analysis does not equate to common sense. By a similar token, for Wittgenstein at least, what people are inclined to say is the starting-point rather than the final word for sober philosophizing (PI §254).

Keeping these different kinds of descriptivism / anti-revisionism apart sheds light on the crucial weakness of D’Oro’s meta-philosophical diagnosis. The meta-philosophical conflict between teleologists and causalists like Jerry Fodor only concerns the merits of descriptive metaphysics / conceptual analysis. It

specifically does *not* include the merits of folk-psychology. Both Davidson and Fodor share a goal with the teleologists, namely that of explaining folk-psychology rather than reforming or abolishing it. Fodor's notorious appeals to "granny" also manifest his proximity to common-sense, at least regarding the existence and centrality of beliefs and desires. And the aforementioned first sentence of ARC makes it abundantly clear that Davidson regards himself as spelling out common sense and its presuppositions rather than as debunking it.

D'Oro herself willy-nilly admits as much. For she takes pains to dispute the claims of causalism to be in line with common sense (e.g. 2012: 217). That, however, is a *philosophical* dispute about our extant mental concepts and common-sense psychological beliefs, rather than a *meta-philosophical* dispute about the role that these concepts and beliefs should play in philosophy. Finally, while D'Oro is right in denying that either common sense or our mental notions imply causalism, she is wrong in maintaining that there is a "common-sensical distinction between actions and events"; instead there is common-sensical distinction, well documented by empirical psychology, between *two kinds of events* – those that result from agency and those that do not. She is also wrong in contending that common sense would accept the Humean claim according to which causal relations boil down to explanatory relations rather than being real relations. On these points Davidson is closer to common sense and our extant conceptual scheme, not than teleologists in general, but than D'Oro's heroes Collingwood and Dray.

6. Naturalism

The conflict between teleologists and causalists is not one between friends and foes of common sense or folk psychology. But these contrasts are not the only meta-philosophical rifts that bear on the reasons and causes debate. More important still is that between naturalism and its discontents, which D'Oro mentions only in passing. So perhaps a slightly different meta-philosophical diagnosis hits the target. One of the widely perceived advantages of causalism is held to be that it is congenial to what friends and foes alike label "the project of naturalising intentional action". As

regards ontology, it seems to dispense with the need to postulate entities beyond the physical realm that nonetheless have the power to cause changes within that realm. As regards epistemology and methodology, it seems to dispense with the need to assume a distinct way of making sense of phenomena that transcends the supremely successful and reasonably well understood causal explanations employed by the natural sciences. Finally, as regards meta-philosophy, it is hospitable to the Quinean idea that the philosophical investigation of agency is part of, or at least continuous with, its empirical investigation by the special sciences.

Because of Quine, many contemporary analytic philosophers, especially in the US, feel almost contractually obliged to profess allegiance to naturalism. This includes staunch defenders of folk-psychology like Fodor. Ironically, however, Davidson may be a heretical exception to this rule, in spirit if not letter. And nowhere more so than in the philosophy of mind and action, his status as an inspiration for physicalism notwithstanding. Admittedly, Davidson denies that there is a clear line between philosophy and science on account of rejecting the analytic/synthetic distinction. This official meta-philosophical naturalism puts him at odds with Wittgenstein. Throughout his career, Wittgenstein sharply distinguished between the “logical”, “grammatical” or “conceptual” problems and investigations of philosophy and the “factual” problems and investigations of empirical science (e.g. TLP 4.122ff.; Z §452). In practice, however, Davidson does not advance a scientific metaphilosophy. If anything, he regards the blurring of the line between philosophy and science as a license to put in perspective the philosophical relevance of neurophysiological findings (1980: 216). Contrasting rhetoric notwithstanding, this is not a million miles away from Wittgenstein’s critical engagement with psychology (see PI II 232; RPP I §§1039, 1093). What is more, Davidson’s repudiation of the analytic/synthetic distinction is incompatible with his idea that the standards of rationality enshrined in his principle of charity are *constitutive* of phenomena like thought, agency and language. Behaviour which completely defies being interpreted as conforming to these standards simply does not qualify as intentional or as linguistic. Davidson himself

was eventually forced to concede that it “cannot be a factual question” whether a creature with thoughts is approximately rational (1985a: 245). This chimes with Wittgenstein’s suggestion that the so-called “laws of thought” are partly constitutive of what we call “reasoning”, “thinking”, and “language” (RFM 80-95, 336).

On occasion, Davidson calls himself a naturalized epistemologist. But what he underwrites is only a “resolutely third person approach to epistemology” (2001: 159, 194). Such an approach is not tied to the transformation of epistemology into physiological psychology that Quine envisaged. In fact, Davidson complains that the latter is inimical to a third-person perspective and that it runs together the physiological causes of our beliefs with their epistemic grounds. Whether these animadversions are compatible with Davidson’s causalism is a moot question (see section 14).

As we have seen, Davidson is precisely *not* a revisionist à la Quine. He is an anti-revisionist, albeit of a variety close to the early rather than later Wittgenstein, in that he regards natural languages and ordinary discourse as structured by logical calculi. This difference is important to Davidson’s theory of action. He derives ontological conclusions about the nature of actions – notably that they are events – from what he takes to be the logical form of action sentences and, more generally, of adverbial modifications (1980: ch. 6). At the same time, Davidson does not rely exclusively on formalisation. His theory of action also features acute observations about established patterns of use. In certain respects, Davidson is a conceptual analyst *malgré lui*. This comes out not just in his astute criticisms of certain versions of agent causation, but also in his famous challenge to anti-causalists, to account for the distinction we standardly draw between *a* reason for A to Φ and *the* reason for which A Φ -ed (1980: 9-10).

This last point also indicates why it is not just inaccurate but also counterproductive to blame the persisting disagreement between causalists and anti-causalists on underlying metaphilosophical divergences. It shifts the bone of contention onto a plane where it becomes even less tractable. As Wittgenstein emphasized (PI §121; see Glock 1996: 243-247), meta-

philosophical disputes are not disputes outside of philosophy; instead, they revolve around “first-order” philosophical issues – often in philosophical logic, philosophy of language and epistemology. Since any stance concerning these issues in turn depends on certain philosophical methods and hence on certain meta-philosophical ideas, we are faced with a potentially vicious circularity. Even leaving that threat aside, meta-philosophical debates tend to draw on a complex syndrome of ideas from diverse areas. That makes them even less amenable to consensual solution than less abstract and fundamental controversies. Tying causalism and anti-causalism *ab initio* to contrasting meta-philosophies renders them more incommensurable than necessary. Viewing the controversy between them in the first instance from a closer angle may bracket underlying divergences that ultimately need to be addressed. But it has the advantage of focusing on the close-quarter exchange of observations and arguments concerning the primary topic of the debate.

7. Shared Ideas in the Philosophy of Mind

There is a context to that debate which is wider than the theory of action yet narrower than meta-philosophy, namely the philosophy of mind. We therefore need to consider whether divergent views about the role of causation in action derive from incompatible assumptions about the nature of the mind. As regards mainstream proponents of the standard story, this is indeed the case. As mentioned above, to them part of the attraction of causalism lies in its prospects for construing both actions and reasons as purely physical phenomena, the former spatially external, the latter spatially internal to the agent.

As regards Davidson himself, however, the picture is more complicated. I already dwelled on his third-person perspective (see also Stoutland 2010: 64). Such a perspective militates against reifying the mind as a kind of entity – be it a Cartesian soul or the brain – and favours linking mental phenomena to agents, their behaviour and abilities. As a result, Davidson’s anomalous monism is much closer to Wittgenstein than commonly recognized. It is “ontological monism coupled with conceptual dualism”. Davidson

tries to reconcile the naturalistic (anti-Platonist and anti-Cartesian) claim that there is no realm beyond the physical with a recognition that mental and semantic discourse is neither reducible to nor replaceable by the idiom of natural science. “There are no such things as minds, but people have mental properties... These properties are constantly changing, and such changes are mental events” (1994: 231; see 1999b: 599).

The key to Davidson’s position is his ontology of events. Like material bodies, events are particulars that can be described in fundamentally different ways. Unlike material bodies, events are dated occurrences with definite durations, like the Olympic Games of 1992 or Brutus’ stabbing of Caesar. They can be described in fundamentally different ways, namely through both mental and physical concepts. *Some* events have true descriptions in psychological terms and are hence mental. But *all* events, including the mental ones, also have true physical descriptions and are hence physical. Thus my coming to believe that *p* has a mental description, e.g. as based on my belief that *q*; but it also has a physical description. The two descriptions pick out *the same* event; indeed, every particular mental event is identical with a particular physical event.

But although there is *token-token* identity between mental and physical events, there can be no *type-type* identity: there are no psychophysical laws that correlate mental events under their mental description (“coming to believe that *p*”) with physical events under their physical description (“ ϕ -fibres firing”). Davidson’s argument in favour of this claim rests on two ideas. First, the principal purpose of mental discourse is to explain behaviour by reference to propositional attitudes such as believing that *p*, desiring \times and intending to Φ . Secondly, this kind of explanation is characterized by two features which are constitutive of mental discourse yet absent from physical discourse – normativity and holism (1980: chs. 11-12).

Mental discourse is normative. Mental states are subject to those principles of rationality which a radical interpreter guided by the principle of charity imputes to the speakers she interprets.

These principles (explored in logic and decision theory) specify what it is reasonable to believe or desire. One such principle is:

(P) If you prefer x to y and y to z , then you should prefer x to z .

Mental discourse is holistic in that propositional attitudes do not occur in isolation. One cannot attribute a propositional attitude to a person independently of attributing other propositional attitudes to her. This is not just an epistemological constraint, but due to the nature of propositional attitudes. The logical connections between propositional attitudes are partly constitutive of their identity: what makes the belief that p the belief it is, is at least in part its interconnection with other beliefs.

Now suppose that there are psychophysical laws which connect preferring x to y and y to z with neural state m and preferring x to z with neural state n . This law should enable us to infer from (P) that if someone is in neural state m he should also be in neural state n . But the only kind of “should” which has a place in physics is one of behaving in line with expectations supported by empirical evidence, not a normative injunction like (P). Equally, this psychophysical law, in conjunction with a physical law according to which anyone in neural state m is also in neural state n , should allow us to predict that whenever someone prefers x to y and y to z he prefers x to z . But because of the holistic nature of thought, the attribution of that last preference may run counter to normative constraints. For example, if we have independent reasons to ascribe a preference of z to x , it lusters the subject with contradictory preferences. Accordingly, physical and mental predicates cannot feature in psychophysical laws because of “the disparate commitments of the mental and the physical schemes” (1980: 222). Two language-games not to be crossed, as Wittgenstein would put it (RFM 117-8).

In Davidson’s view, “the mental is not an ontological but a conceptual category” (2004: 46). His conceptual dualism and refusal to reify mental notions includes reasons for actions.

Beliefs, desires and intentions belong to no ontology When we ascribe attitudes [like believing, desiring and intending something], we are using the mental vocabulary to describe people. Beliefs and

intentions are not little entities lodged in the brain. (1999b: 654; see also 2001: 60)

According to Davidson, there is an analogy between saying that a subject thinks that *p* and that a subject weighs *n* kg (2001: ch. 4; on the measurement analogy see also Beckermann 1996; Matthews 2007.) When we ascribe a weight to a person, we do not ascribe to them a genuine relation to an abstract object. Rather, we ascribe to the person a relation to other material objects, for instance that it would be in balance with 60 litres of water. *Mutatis mutandis* for the case of belief. In ascribing a belief to a person, we ultimately describe and explain their actual or possible behaviour. We place the subject not in a relation to a genuine object, but in the context of a system of describing and explaining the subject's behaviour and behavioural capacities. In the final analysis, so-called propositional attitudes are to be elucidated in terms of what subjects think or say, or, more accurately still, *could* think or say. In sum: there is a striking though surprisingly unappreciated proximity between Davidson's approach on the one hand, and Wittgenstein's insistence on the conceptual connection between the mental and behavioural capacities and his idea that mental predicates apply first, foremost and literally to human beings rather than to souls or parts of bodies, the brain included (PI §281).

8. Davidson and analytic hermeneutics

The gap between Wittgenstein and Davidson in the philosophy of mind is too narrow to render Wittgensteinian teleology and Davidsonian causalism incommensurable. Even in the theory of action, the two are closer than commonly recognized, even on a fairly orthodox construal.⁵ In fact, Davidson is a *bona fide* member of analytic hermeneutics. This is no coincidence, given his intellectual biography (see 1999a: 12-27). It derives partly from his interest in literature and history as an undergraduate and from his study of ancient philosophy – notably Plato and Aristotle – as a graduate student. In the course of writing his PhD dissertation on

⁵ For an unorthodox interpretation that assimilates the two even more see Stoutland 2007.

the *Philebus* he came to appreciate an essay by Gadamer. Much later Davidson devoted an article to this essay, and there are brief yet respectful criticisms of Gadamer in his last writings (2005: 181, 252; ch. 18). More importantly, Davidson knew most of the British representatives of analytic hermeneutics, including Ryle, Anscombe, Hampshire and Kenny. He invited several of them to Stanford in the 1950s and spent a year at Oxford in 1973. Qualms about “small red books” notwithstanding, he admired much in their work. On the back cover of the 2000 reissue he lauded Anscombe’s *Intention* as “the most important treatment of action since Aristotle”. And he would have readily conceded that it was Wittgenstein, Ryle and their followers who re-established the theory of action as a flourishing discipline.

As regards doctrinal matters, Davidson combines the hermeneutic idea that actions are explained by the reasons for which agents perform them with a causal theory, according to which actions are caused by those reasons – or the events of those reasons coming into existence – and explained in a causal fashion. Both actions and the reasons that explain them are among those events that have both a physical *and* a mental description. Under their physical descriptions, the action can be seen to be *caused by its reason*. At the same time, there is *no* strict causal law which relates action and reason as described in mental terms. For which mental descriptions hold for these events depends on considerations of rationality that are qualitatively distinct from physical considerations.

As a result of his conceptual dualism, Davidson also endorses the methodological dualism of hermeneutics: “there is an irreducible difference between psychological explanations that involve the propositional attitudes and explanations in sciences like physics and physiology” (2004: 101). He also subscribes to Collingwood’s view that “the methodology of history (or, for that matter, any of the social sciences that treat individual social behaviour) differs markedly from the methodology of the natural sciences” (2005: 282).

In summary, for Davidson, human beings and their actions are physical; yet they cannot be properly understood by physics, but

only by appeal to reasons and rationality. He shares Quine's predilection for a monistic ontology, but not for a scientific methodology. In this way, he seeks to pay the ontological homage that philosophy owes to modern science, without denying the distinctive character of human beings. However, the devil lies in the details of this combination. There Davidson diverges from Wittgenstein, but even more so from some of the small red books. It is to them and to Davidson's riposte that we must now turn.

9. The Teleological Arguments against Causalism

ARC explicitly defends causalism against five, partly interconnected, arguments

- a. States vs. events: Reasons consist of attitudes and beliefs, and these are states or dispositions; yet only events can be efficient causes.
 - b. Logical connection: The connection between an action and the reason for which it is done is logical or conceptual rather than empirical. The two cannot be stated independently of one another, by contrast to the relation of a causal relation as conceived in the Humean paradigm.
 - c. Absence of causal laws: There are no causal laws connecting intention and action. By contrast, an event A can cause an event B only if events of type A are connected with events of type B by a general *law*.
 - d. First-person authority: By and large, agents have first-person authority about the reasons for which they acted. I.e. what they sincerely claim to be their reason *is* what we call their reason. By contrast, their statements of the causes of their actions enjoy no special authority; they are no more than fallible hypotheses.
 - e. The problem of the missing agent: If reasons were inner events causing the action, the agent would be a helpless victim rather than the author of the action.
- (b) – (e) appear in writings by Wittgensteinian teleologists. They also have resonances in Wittgenstein's own work. This is not to say that Wittgenstein subscribed or would have subscribed to all of

them. He condoned versions of (c) and (d), yet in a way that is not as irreconcilable with Davidson as one might suspect. He did not advance (b); indeed, his account of logical connections implies that (b) is fallacious. I shall disregard (e), which is related to Wittgenstein's Schopenhauerian animadversions against the empiricist idea that "willing too is merely an experience" (see Glock 1999; Candlish 2001). Instead, I shall consider (b) – (d) in the light of Wittgenstein's own reflections on agency and causation. In the final sections I turn to the connection between (a) and a Wittgensteinian objection not mentioned in ARC, namely the distinction between causal explanation and rational justification.

10. Davidson's Wittgensteinian Critique of the Logical Connection Argument

The logical connection argument received its emblematic – though not exemplary – formulation by Melden. According to Melden, a lesson from Hume is that a cause "must be distinct from the alleged effect". Yet the connection between the reason for an action and the action is *logical*: the reason for Φ -ing is identified as the reason it is through being a reason for Φ -ing, i.e. through the action for which it is a reason (1961: 52-3). Melden illustrates his objection through the example of a driver raising his arm in order to signal that he is taking a turn. "If ... the motive were some event either concurrent with or antecedent to the action of raising the arm, there would need to be a logically necessary connection between two distinct events – the alleged motive and the action, however it is described. This is impossible if the sequence motive \rightarrow action is a causal relation" (1961: 89).

Davidson renders the argument as follows: "Since a reason makes an action intelligible by redescribing it, we do not have two events, but only one under different descriptions, *while* causal relations (in the Humean sense) demand two distinct events" (1980: 13-4). He raises two main objections.

First, a cause differs from its effect; but a reason also differs from the action it rationalizes. For one thing, they differ as regards their *categorial status*: "reasons, being beliefs and attitudes, are

certainly not identical with actions”, which are events (1980:10). For another, they differ as *individual phenomena*. It is perfectly possible, for instance, to want or intend to turn on a light without actually doing so.⁶

Both responses are compelling. However, they leave intact a different version of the logical connection argument. von Wright accommodates the second point by granting that the logical connection between the reason and the action obtains only *ex post actu* (1971: 116-7, see 93-4, 107). At the same time, he insists that the connection between the premises and the conclusion of a practical inference is conceptual/logical rather than empirical/causal. If an agent wants or intends to Φ at t , she will do so, unless she is somehow prevented from Φ -ing at t . Anscombe puts the point as follows: “the primitive sign of wanting is *trying to get*” (1957: §36). In Wittgenstein’s own terms: even though one can want or intend to Φ without actually Φ -ing, Φ -ing or trying to Φ is a *criterion* rather than a symptom of wanting or intending to Φ . The terms “wanting to Φ ” and “intending to Φ ” would not mean what they do, unless A ’s failing to try to Φ under propitious circumstances constituted evidence, however defeasible, against “ A wants/intends to Φ ”.⁷

This marks an important difference between the *way we standardly talk about* the relation between reason and action on the one hand, and the way we talk about the relation between relations of efficient causation. Yet for reasons mentioned in section 3 it does not rule out that reasons *are* causes (causalism about reasons), and hence it is also compatible with a causalism about explanation that insists that intentional explanations refer to efficient causes, albeit in a conceptually distinct manner.

⁶ This riposte is standardly imputed to ARC (e.g. Davis 2010: 37). But it is only implicit on pp. 14-5.

⁷ For the pertinent features of Wittgenstein’s account of criteria for mental concepts in general see Glock 2001: 9-12. Wittgenstein did not apply this account to the case of wanting in the way advocated here. Waismann (1983: 12-5, 36-40) does not employ the notion of criteria, yet comes close in other respects, through his modifications of Schopenhauer’s dictum *Der Wille ist die Tat* (“the will is the deed”).

This is where Davidson's second response comes in (1980: 14-15). The logical connection argument assumes that causal relations are *factual* rather than *logical*. But consider the synthetic truth

(1) A caused B

Since the cause of B = A, we can move to the analytic

(2) The cause of B caused B.⁸

The *logical status* of a statement linking events as analytic or synthetic depends not on the events themselves, but on *how they are described*. The same holds for the *difference in explanatory value* between (1) and (2). Two events can be logically related under *one pair of descriptions*, while at the same time being factually related under *another pair*. Speaking less loosely, one can describe A and B correctly in different ways, such that some descriptions of A are logically related to some descriptions of B, while other descriptions of A are not logically related to other descriptions of B. As a result, the fact that two events *can* be described in a way that relates them logically is no bar to them also being related as cause and effect. Indeed, redescribing events in terms of their causes is as much part of ordinary discourse as redescribing actions in terms of their reasons. "Suppose that someone was injured. We could redescribe this event 'in terms of a cause' by saying he was burned" (1963: 10).

Both of Davidson's responses employ conceptual analysis to good effect. Furthermore, far from being incompatible with Wittgenstein, the second response starts out from a *bona fide* Wittgensteinian idea (Glock 1996: 75-6, 129-35, 198-202). Davidson stresses that logical relations are not *de re* but *de dicto*, i.e. due to the way we describe things, rather than to mind- or language-independent necessities. The idea can be classified as *conventionalist*, and in Davidson it combines Quinean and Wittgensteinian influences. Although Quine rejected *de re*

⁸ The cause of B caused B is not analytic, because there might not have been any cause of B, Miguel Hoeltje demurred. But one can reformulate (2) as "If anything caused B, then the cause of B caused B". And that is analytic. It exemplifies a valid pattern: "If anything V-ed X, the thing that V-ed X V-ed X". Furthermore, in the spirit of Strawson one should insist that talk of "the cause of B" *presupposes* that B had a cause.

necessities, he did not adumbrate *de dicto* necessities, instead rejecting necessity altogether. At least in ARC Davidson accepts *de dicto* necessities. He could have inherited that idea from Carnap, but the latter in turn got it from Wittgenstein. In any event, it is more probable that Davidson imbibed it from Anscombe's emphasis on actions having intentional properties "under a description" (1959: 11-12). It allows Davidson to accept that the way we describe reasons and actions in rationalizations is unique, while nonetheless holding that these rationalizations refer to efficient causes and their effects.

11. The Absence of Causal Laws

(b) maintains that general laws are implied by causal explanations, yet not by intentional explanations. It concludes not just that rationalisations differ from causal explanations (contrary to causalism about explanation), but also that reasons are not causes (contrary to causalism about reasons). While Davidson accepts the first half of the premise, he rejects the second and hence the conclusion. Just as his later anomalous monism revolves around the idea that there are no psychophysical laws, ARC denies that there are *psychological* laws connecting beliefs, desires etc. with actions. There are "rough laws", "generalisations connecting reasons and actions", yet these cannot be sharpened into strict predictive laws. However, this "does not inhibit valid causal explanation, or few causal explanations could be made. I am certain the window broke because it was struck by a rock – I saw it all happen; but I am not (is anyone?) in command of laws on the basis of which I can predict what blows will break which windows" (1980: 15-16). "Singular causal statements" like "A causes B" do not imply a "particular law involving the predicates used in the descriptions 'A' and 'B'", i.e. a law of the form "Events of type A are followed by events of type B". At this point, Davidson departs from Hume. But he has the later Wittgenstein on his side.

Like Davidson, Wittgenstein accepts one feature of the Humean view: causal relations are *external*, i.e. obtain between logically independent events (PI §220; Z §296). And *like* Davidson, he rejects another feature, the nomological account of causal

explanations: not all causal explanations are underpinned by general connections between types of events. There is an irreducible variety of *prototypes* of causal connections, including: (i) impact (collision of billiard-balls); (ii) traction (pulling a string); (iii) mechanisms like clocks, which combine (i) and (ii); (iv) human reactions to sensations or emotions (being hit on the head or frightened by someone's facial expression); (v) statements which are based on observing regular successions of events. Since Wittgenstein stresses both the variety of cases and the fact that we use the same word he arguably regards “cause” as a family resemblance concept. He denies not only that the Humean paradigm (v) is the only prototype of causation, but also that it is the fundamental one. And he does so in a way that is very close to the passage from Davidson just quoted. The “cause-effect language game” of everyday life is rooted not in observation or experimentation revealing general laws, but in a *practice*, which in turn is based on certain primitive *reactions*. For example, we react to a painful blow by pointing to someone and saying “He did it” (CE 409-10, 416-7, 420, 433). The crucial lesson for us: the basic cases of causal explanations are those of *singular causation*, which do not require general laws.⁹

A difference between Wittgenstein and Davidson emerges elsewhere. Davidson accepts a “weaker” version of Hume’s nomological account of causation. (1) does entail “that there exists a causal law instantiated by *some true descriptions* of A and B”. This weaker version “suits rationalizations equally well”. If a Φ s because a desires that p and believes that Φ -ing will bring it about that p , then there exists a strict causal law at the neurophysiological level connecting the onslaught in a of the pro-attitude / belief causation and a’s bodily movement (1980: 16-7).

This is the message of Davidson’s anomalous monism, but one to which Wittgenstein would be deaf. He is committed to denying that a causal relation can obtain only if there is some causal law –

⁹ Tripodi points out that Anscombe (1971) is close to Davidson on this score, since she accepts not just *non-necessitating* causes but also *singular* causation. Both ideas go back to Wittgenstein (Glock 1996: 72-4).

however unknown or even unknowable – connecting A and B. This is implied by his animadversions against the idea that whenever an effect occurs in one case but not in an apparently similar case, there must be some relevant further differences, even though for reasons of principle we may never be in a position to identify it. In the case of two apparently identical plant-seeds which produce different kinds of plants, Wittgenstein avers, there *need not be* any difference in the seeds underlying these different dispositions. We could and should treat the origin of these seeds not just as the basis for a prediction (“Seeds from a type A-plant *will* produce type A-plants”), but add “... *because* they are from type A-plants”. Wittgenstein gives the impression that the origin could furnish not just a *bona fide* causal explanation, which is in line with Davidson’s aforementioned passage, but the ultimate explanation which captures all there is to the causal relation. And he writes: “If this upsets our concepts of causation, it is high time they were upset” (Z §§608-10; see CE pp. 410-11, 433-4; LPP 100-1).

This line of reasoning is not compelling. By Wittgenstein’s own lights, philosophy has no business “upsetting”, i.e. revising our notions of causation. In this passage, he appears to be less of a descriptivist and conceptual analyst than Davidson (once more, contrary to d’Oro’s meta-philosophical diagnosis). Furthermore, even *if* there were nothing *unintelligible* in supposing that there is no structural difference between the two seeds, accepting *phenomenal* properties (concerning the origin of the seeds) as ultimate causes would mean abandoning a highly successful principle of the physical sciences. It would be on a par with accepting that astrological explanations capture the ultimate factors shaping our destinies, provided that they are backed by statistical evidence. But even if it should turn out, for example, that people born between 21 January and 18 February have an above average IQ, it would be irrational to insist that this fact was brought about by the fact that these people were born under the sign of Aquarius.

12. First-Person Authority

Anti-causalist argument (c) runs: an agent knows of her own reasons for acting infallibly, without induction or observation; yet

knowledge of ordinary causal relations is fallible; hence intentional explanations differ from causal explanations and reasons are not causes. ARC accepts the first premise, with qualifications. Davidson points out that agents can be wrong about the reasons for which they acted, in cases in which they have more than one reason. But like the Wittgensteinian teleologists he denies that “in general it makes sense to ask you how you know what your reasons were or to ask for evidence” (18). At the same time he rejects the conclusions, since he does not accept the second premise. Induction or observation is “not the only way of knowing” that a causal law exists and hence of knowing that a causal relation obtains.

That final conclusion, at least, is once more in line with Wittgenstein, for reasons connected with those concerning singular causation sketched above. According to Hume we can never directly observe a causal connection, but only a succession of events; consequently our causal statements must be based on observing a regular sequence of parallel events and are always provisional, subject to refutation by subsequent observations. Wittgenstein follows Russell in holding that there are causal relations which we know *immediately*, while rejecting the idea that this is based on intuition (CE 409, 431; LC 22). Recognizing the most basic forms of causation, especially those involving direct physical contact (i)-(iv), does not depend on observing constant regularities or on experiments. We directly observe one thing acting upon another, and know the cause immediately, though not infallibly. Both are paradigmatic cases of what we call a cause, and constitutive of the idea of a causal nexus (CE 408-10, 416, 420, 433).

Acknowledging the possibility of non-inductive, non-observational knowledge of singular causal connections is compatible with insisting that as regards *her intentional actions*, an agent does not have the kind of authoritative knowledge of their causes as she does of their reasons. It is noteworthy, however, that Anscombe, in applying Wittgenstein’s idea of immediate causal knowledge to the case of agency, accepts that agents can have immediate knowledge of the causes of certain things they do. She

reckons with such a thing as a “mental cause”, “what someone would describe if he were asked the specific question: what produced this action or thought or feeling on your part: what did you see or hear or feel, or what ideas or images cropped up in your mind, and led up to it?”. For instance, when asked “Why did you knock the cup off the table?” one might respond “I saw such-and-such and it *made me jump*”. Anscombe rightly notes that we know these things authoritatively, “without observation” (1957: 16-18).

On the one hand, Anscombe wants to distinguish such mental causes from motives and intentions. On the other hand, she writes:

Intentional actions are a sub-class of the events in a man’s history which are known to him *not* because he observes them. In this wider class is included one type of involuntary actions, which is marked off by the fact that mental causality is excluded from it; and mental causality is itself characterized by being known without observation. But intentional actions are not marked off just by being subject to mental causality, since there are involuntary actions from which mental causality is not excluded (1957: 24, cp. 18-9).

This passage strongly suggests that intentional actions are a sub-class of actions brought about by mental causality; otherwise it would have to read somewhat as follows: “intentional actions are marked off by not being subject to mere mental causation”. If this reading is correct, Anscombe was a Davidsonian *avant la lettre*.

13. Explanation vs. Justification

There is one Wittgensteinian argument that Davidson does not mention in ARC, presumably because it does not loom large in the writings of the Wittgensteinian teleologists. This is surprising, since the objection is rather straightforward. It concerns the contrast between justification and explanation. A mental state or process – e.g. of believing or coming to believe that *p* – may causally explain an action, but it cannot provide a justification for that action.

The contrast between explanation and justification was originally pinpointed in Kant’s distinction between *quaestio facti* and *quaestio iuris* and the ensuing Neo-Kantian distinction between “genesis” and “validity”. It has fuelled a pervasive, if largely implicit, suspicion of the so-called “genetic fallacy”, the mistake of

deducing claims about the *validity* of a statement or the *content* of a concept from information about the causes of its emergence (see Glock 2014; Glock 2008: 100-103). My conjecture is that Frege imbibed it from Neo-Kantianism and passed it on to Wittgenstein and Waismann. In Frege, this Kantian anti-geneticism was combined with a distinction that concurrently emerged within the phenomenological tradition, namely between a mental act and its object, or, more appositely, between a mental state or process like that of believing or desiring and its content – that which is believed or desired. In many cases, at least, one should further distinguish that content of the believing from its object(s) – what the believing is about. When A believes that the cat is on the mat, *that the cat is on the mat* is the content of A's believing, whereas *the cat* and *the mat* are its objects.

In his attempt from the 1930s to summarize Wittgenstein's position, Waismann applied the explanation/justification dichotomy to the theory of action – the activity of following a rule – by availing himself of a specific case of that distinction.

The attending to the rule can indeed be the cause for the rule being followed. ... [But] the cause of an action can never be referred to, to justify the action. I may justify a calculation by appealing to the laws of arithmetic, but not by appealing to my attending to these laws. The one is a justification, the other a causal explanation. (Waismann 1965: 123)

What could be invoked as a *reason*, as a justification of that person's action, is not *the attending to the rule*, but the rule attended to. More generally, the reason for an agent's action is not the believing or desiring, but the *content of the believing or desiring*. The reason for A's Φ -ing is not A's believing or desiring that p, but just *that p*. And that is not a mental cause of A's Φ -ing, simply because it not a *mental* phenomenon at all.

Wittgenstein himself employed a distinction between the causes and the contents of believing and desiring in his discussions of intentionality in general and of aesthetic judgments in particular. He also linked this distinction to the idea that the reasons for a belief, unlike its causes, are what is believed. "When I am asked for

a reason for the belief, what is expected, as part of the answer, is *what I believe*” (AWL 28).¹⁰

Tripodi (2014) detects this point in Anscombe, though only *in nuce*.¹¹ Be that as it may, it has become central to the current debate. Contrary to a prevailing impression, the causalist consensus established by ARC has recently been challenged. Its Achilles-heel is precisely the claim that reasons are mental phenomena rather than objective facts or presumed facts. At present, two contrasting positions about reasons for action confront each other. First, there is an account which can be labelled psychologistic, psychological, internal or subjectivist. According to this subjectivist conception,

(S) *A* Φs for a reason iff *A*'s Φ-ing is to be explained by reference to mental states of *A* (beliefs and “pro-attitudes” like desires and intentions).

Subjectivism insists that the beliefs and desires invoked in intentional explanations are *mental states* of the agent – states of desiring or wanting something and states of believing something.

By contrast, according to a *revisionist* position that is currently gaining in momentum, the reasons for which agents act are not mental states of the agent herself. Instead, they are objective *facts* or states of affairs, facts which, save for special cases, concern the agent's environment rather than her own mind. In so far as the reasons for which an agent acts can be said to be beliefs and desires at all, they are not subjective states of believing or desiring, but *what is believed* or *what is desired*. An agent *A* acts for a reason if *A* acts on account not of her own beliefs, but on account of facts, on account of how things are (from *A*'s perspective). Accordingly,

(O) *A* Φs for a reason iff *A* acts in the light of reasons, that is, in the light of facts (as *A* sees them).

¹⁰ I am grateful to Matthieu Queloz. His MA thesis reminded me of the first point (discussed in Glock 1996: 185-6, 33-4) and alerted me to the striking passage from AWL. Queloz also cites other passages, but they are either far less explicit or variants of the passage from Waismann 1965 hailing from VW.

¹¹ As regards “backward looking motives” like revenge, Anscombe writes: “something that has happened ... is given as the ground of an action” (1957, § 13). For instance, the reason *A* gives for an act of revenge against *B* is the fact that *B* killed *A*'s brother.

Subjectivism goes back at least to Hume, and the “standard story” closely associates its present manifestations with Davidson. However, that attribution is far from obvious. To see why, it helps to look at the anti-causalist argument (a) considered in ARC.

14. Reasons, Beliefs and “Onslaughts”

Although this argument is congenial to Wittgensteinian teleologists, it is absent from their writings. But this is unsurprising, since it is an objection Davidson raises against *his own version* of causalism. For Davidson, A’s reason for Φ -ing is roughly speaking A’s mental state of believing that p will lead to q, in combination with her mental state of entertaining a “pro-attitude” or desiring that q (Glock 2013: 10). Accordingly, reasons are pro-attitudes and beliefs. These are states or dispositions, categorically distinct not just from actions (see sct. 10), but also from events. Only events, however, can be efficient causes. Therefore, the objection concludes, reasons cannot be efficient causes (contra causalism about reasons) and intentional explanations make no reference to causes (contra causalism about explanations).

Davidson considers two responses. One is to regard reasons not as causes *strictu sensu* but as “causal conditions”. The other is to insist that the causes of actions consist in the “onslaught” of the primary reason, i.e. the coming into existence of the combination of belief and pro-attitude. These responses are obviously compatible. Yet even combined they will not reinstate Davidson’s official doctrine, enshrined in the standard story. After all, he propounds the slogan: reasons are causes (ARC 12, 4). But his responses only entitle him to the claim that reasons are causal conditions of action and that their onslaughts are the causes of action.

Even this modified version of Davidson’s doctrine is subjectivist. Mental phenomena play a *dual role*: They both *cause* and *rationalize* the action. According to some recent commentators, however, Davidson’s position, rightly understood, actually converges on objectivism. The item that causes *differs* from the item that rationalizes. It is the occurrence of the attitude (“onslaught”), i.e. a mental event, which causes the action. The causal relation

holds between two occurrences, mental event and bodily movement. It is the content of the attitude which rationalizes the action. The relation of rationalization holds between the propositional content of the mental attitude and the description of the action (e.g. Keil 2012).

All very well, since objectivism is not just closer to Wittgenstein than subjectivism, but also to the truth (Glock 2012), except that this modification of Davidson's theory simply amounts to *abandoning* causalism about *reasons*. It concedes that reasons simply are *neither* mental *causes* – i.e. events of onslaughts – *nor* even mental causal *conditions* – i.e. states of believing and desiring brought about by these onslaughts.

This concession has implications not just for Davidson's relation to Wittgenstein, but also for his famous accusation that Quine's theory of meaning is guilty of a "third dogma of empiricism", the dualism of conceptual scheme and empirical content. According to Quine, the stimulus-meaning of simple sentences consists of the set of neural stimulations that would prompt assent to that sentence. Davidson complains that this account posits "epistemic intermediaries" to intervene between the world and our beliefs and utterances (1984: ch. 13; 2001: chs. 3, 10). He accuses Quine of succumbing to the empiricist "myth of the given", the idea that pre-conceptual sensory stimulations provide the foundations of knowledge and of meaning.

Davidson denies that neural stimulations can play either the epistemic role of providing the evidence for our beliefs or the semantic role of determining the meaning of our sentences. Neural firings feature in the causal chain between objects and events on the one hand, beliefs and assent on the other. But to think of them as evidence is to confuse the causes of our beliefs with the "reasons" or "justification" on which they rest. Neural events cannot stand in relations of either logical or probabilistic support (2001: 141-4; 2004: 69-70).

Prima facie, this line of criticism is compatible with Davidson's thesis that reasons are causes. The latter only implies that all reasons are causes, *not* that all causes are reasons, thereby allowing that neural stimulations can be causes without being reasons.

However, Davidson's central tenet is that events can be described in different idioms. He also insists that logical relations are *de dicto*, i.e. due to the way we describe things, whereas causal relations hold between events no matter how described. Therefore Davidson must allow that there are descriptions of neural events under which they stand in logical or probabilistic relations to beliefs. The retinal stimulation that causes me to see a sign-post ahead marks the onslaught of my belief that there is a sign-post ahead, yet it also stands in a logical relation to that belief. An obvious remedy is to distinguish between the neural cause of my believing and the reason, which is neither the mental state of believing nor its neural cause, but rather what is believed – an objective state of affairs, namely that there is a sign-post ahead. This remedy is suggested not merely by the interpretation just mentioned, but also by some of Davidson's own remarks (e.g. 1980: xi). However, it implies that there is a difference between the item that causes the believing – the neural event – and the item that justifies it – its content. Assuming a strict identity of reasons and causes, Davidson is at best entitled to insist against Quine on a distinction between *causal explanation* and *epistemic justification*, as when he writes “a causal explanation of a belief does not show how or why the belief is justified”. He is not entitled to invoke an *ontological* distinction, e.g. when he inveighs against “transmuting a cause into a reason” or writes “even if our *reasons* for our beliefs are always other beliefs, the *causes* sometimes lie elsewhere” (2001: 143, 169).

15. Causalism about Explanation and Davidson's Challenge

What is left of the Davidsonian position as it features in the standard story is causalism about intentional explanations. The reasons for which an agent acts are facts, as she sees them. But the explanation of her actions refer to her coming to believe these facts or presumed facts.

As regards the tenability of causalism about explanation, the seminal contribution of ARC is a challenge to teleologists, namely to explain the difference between a reason A had for Φ -ing and the reason for which A actually Φ -ed. It is the challenge to explain the

“mysterious” special connection between *the* reason and the action without invoking the notion of a cause (1980: 11). In my view one can distinguish three responses to this challenge.

- *Third Generation Causatism*: shoring up the causalist picture, notably through modifying (widening) the notion of causation.
- *Patternalism*: what distinguishes the reason from a reason is the role that the reason (propositional content) plays and has played in the agent’s mental and public biography – whether it fits into the *pattern* of her thoughts, utterances, and deeds.
- *Deliberationism*: what distinguishes a reason from the reason is that only the reason weighed in A’s deliberations, either *ante* or *post factum*.

Patternalism and, to a lesser extent deliberationism, are inspired by Wittgenstein. Are any of these responses capable of meeting Davidson’s challenge, either singly or in combination? That question must be left for another occasion.¹²

List of References

- Alvarez, M. 2007. “The Causalist/Anti-causalist Debate”. In: A. Leist (ed.), *Action in Context*. Berlin: de Gruyter, pp. 103-123.
- Candlish, S., 2001. “The Will”. In: H-J. Glock (ed.), *Wittgenstein – A Critical Reader*. Oxford: Blackwell, pp. 156-173.
- Candlish, S. and Damjanovic, N., 2010. “Reasons, actions and the will: the fall and rise of causalism”. In: M. Beaney (ed.), *The Oxford Handbook of the History of Analytic Philosophy*. Oxford: Oxford University Press, pp. 689-708.
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford: Clarendon Press.
- Davidson, D. 1984. *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.
- Davidson, D. 1985a. “Replies to Essays”. In: B. Vermazen and M. Hintikka (eds.), *Essays on Davidson: Actions and Events*. Oxford: Clarendon Press.

¹² For comments, discussions or editorial help I should like to thank Meret Polzer, Angela Ray, Constantine Sandis, Paolo Tripodi, as well as participants in my graduate seminar in the theory of action and my co-symposiasts at the conference “50 Years of Actions, Reasons and Causes” (Mülheim 2013).

- Davidson, D. 1985b. "Reply to Quine on Events". In: E. LePore and B. McLaughlin (eds.), *Actions and Events: Perspectives on the Philosophy of Donald Davidson*. Oxford: Blackwell, pp. 172-176.
- Davidson, D. 1990a. "The Structure and Content of Truth", *Journal of Philosophy* 87, pp. 279-328.
- Davidson, D. 1994. "Donald Davidson". In: S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*. Oxford: Blackwell, pp. 231-36.
- Davidson, D. 1999a. "Intellectual Autobiography". In: L.E. Hahn (ed.), *The Philosophy of Donald Davidson*. Library of Living Philosophers, vol. XXVII. Chicago: Open Court, pp. 3-79.
- Davidson, D. 1999b. "Replies to Critics". In: L.E. Hahn (ed.), *The Philosophy of Donald Davidson*. Library of Living Philosophers, vol. XXVII, Chicago: Open Court.
- Davidson, D. 2001. *Subjective, Intersubjective, Objective*. Oxford: Clarendon Press.
- Davidson, D. 2004. *Problems of Rationality*. Oxford: Clarendon Press.
- Davidson, D. 2005. *Truth, Language, and History*. Oxford: Clarendon Press.
- Davis, W.A., 2010. "The Causal Theory of Action". In: O'Connor and Sandis (eds.), *A Companion to the Philosophy of Action*, pp. 32-39.
- D'Oro G., 2012. "Reasons and Causes: The Philosophical Battle and the Meta-philosophical War". *Australasian Journal of Philosophy*, pp. 207-221.
- Frascolla, P., 2010. D. Marconi and A. Voltolini (eds.), *Wittgenstein: Mind, Meaning and Metaphilosophy*. London: Palgrave.
- Glock, H.J., 1996a. *A Wittgenstein Dictionary*. Oxford: Blackwell.
- Glock, H.J., 1996b. "Abusing Use". *Dialectica*, pp. 205-223.
- Glock, H.J., 1999. "Schopenhauer and Wittgenstein: Representation as Language and Will". In: C. Janaway (ed.), *The Cambridge Companion to Schopenhauer*. Cambridge and New York: Cambridge University Press, pp. 422-458.
- Glock, H.J., 2000. "Animals, Thoughts and Concepts", *Synthese* 123, pp. 35-64.
- Glock, H.J., 2001. "Wittgenstein and Quine: Mind, Language and Behaviour". In: S. Schroeder (ed.), *Wittgenstein and Contemporary Philosophy of Mind*. Basingstoke: Palgrave, pp. 3-23.
- Glock, H.J., 2003a. *Quine and Davidson on Language, Thought and Reality*. Cambridge: Cambridge University Press.
- Glock, H.J., 2008. *What is Analytic Philosophy?* Cambridge: Cambridge University Press.
- Glock, H.J., 2010. "Does Language require Conventions?". In: P. Frascolla, D. Marconi and A. Voltolini, 2010, pp. 85-112.
- Glock, H.J., 2013. "Quine and Davidson". In: K. Ludwig and E. LePore, (eds.), *The Blackwell Companion to Davidson*. New York: Wiley, pp. 567-587.

- Glock, H.J., 2014. "Neo-Kantianism and Analytic Philosophy". In: N. de Warren and A. Staiti (eds.), *The Legacy of Neo-Kantianism*. Cambridge: Cambridge University Press, forthcoming.
- Grice, H.P., 1989. *Studies in the Way of Words*. Cambridge/Mass.: Harvard University Press.
- Hacker, P.M.S., 1996. *Wittgenstein's Place in Twentieth Century Analytic Philosophy*. Oxford: Blackwell.
- Hempel, C.G., 1942. "The Function of General Laws in History". *Journal of Philosophy* 39, pp. 35-48.
- Hornsby, J., 1980. *Actions*. London: Routledge Kegan Paul.
- Matthews, R.J., 2007. *The Measure of Mind*. Oxford: Oxford University Press.
- Melden, A. I., 1961. *Free Action*. London: Routledge & Kegan Paul.
- Rhees, R., 1965. "Some Developments in Wittgenstein's View of Ethics". *Philosophical Review* 74, pp. 17-26.
- Ryle, G. 1949. *The Concept of Mind*. London: Penguin, 1980.
- Smith, M., 2004. "The Structure of Orthonomy". In: J. Hyman and H. Steward (eds.), *Agency and Action*. Cambridge: Cambridge University Press, pp. 165-94.
- Stoutland, F., 2010a. "Reasons and Causes" In: P. Frascaola, D. Marconi, A. Voltolini (eds.), *Wittgenstein: Mind, Meaning and Metaphilosophy*, pp. 46-66.
- Stoutland, F., 2010b. "Von Wright". In: C. Sandis and T. O'Connor (eds.), *A Companion to the Philosophy of Action*, pp. 589-597.
- Sandis, C., 2015. "Small Red Book Historicism", forthcoming in *The Journal of the Philosophy of History*.
- Sandis, C., and O'Connor, T., (eds.) 2010, *A Companion to the Philosophy of Action*. Oxford: Wiley-Blackwell.
- Tripodi, P., 2014. "Davidson and the Wittgensteinians on Reasons and Causes", forthcoming in: D. Moyal-Sharrock, V. Munz, A. Coliva (eds.), *Language, Mind and Action. Proceedings of the 36th Wittgenstein Symposium in Kirchberg am Wechsel*. Berlin: Walter de Gruyter.
- Waismann, F., 1965. *The Principles of Linguistic Philosophy*. R. Harré, ed. 1965. London: Macmillan.
- Waismann, F., 1983. *Wille und Motiv*. J. Schulte, ed. 1983. Stuttgart: Reclam.
- Wittgenstein, L., 1980. CV: *Culture and Value*, G. H. von Wright (ed.), P. Winch (trans.). Oxford: Blackwell.
- Von Wright, G.H., 1963. *Norm and Action*. London: Routledge & Kegan Paul.
- Von Wright, G.H., 1971. *Explanation and Understanding*. Ithaca/N.Y.: Cornell University Press.

Biographical Note

Hans-Johann Glock is Professor of Philosophy at the University of Zurich (Switzerland), and Visiting Professor at the University of Reading (UK). He previously held positions at Oxford and Reading, as well as visiting professorships and research fellowships in Canada, Germany and South Africa. He is the author of *A Wittgenstein Dictionary* (Blackwell 1996), *Quine and Davidson on Language, Thought and Reality* (CUP 2003), *La mente de los animales* (KRK 2009) and *What is Analytic Philosophy?* (CUP 2008). He has edited *The Rise of Analytic Philosophy* (Blackwell 1997), *Wittgenstein: a Critical Reader* (Blackwell 2001) and *Strawson and Kant* (OUP 2003), and co-edited (with Robert L. Arrington) *Wittgenstein's Philosophical Investigations* (Routledge 1991), *Wittgenstein and Quine* (Routledge 1996) and (with John Hyman) *Wittgenstein and Analytic Philosophy* (OUP 2009). He has published numerous articles on the philosophy of language, the philosophy of mind, the history of analytic philosophy and Wittgenstein. He is the recipient of a Humboldt Foundation Research Prize for 2014-15, and hopes to use the prize to finish a book on animal minds.