A Modest Proposal for Interpreting Structural Explanations
Mariam Thalos

ABSTRACT
Social sciences face a well-known problem, which is an instance of a general problem faced as well by psychological and biological sciences: the problem of establishing their legitimate existence alongside physics. This, as will become clear, is a problem in metaphysics. I will show how a new account of structural explanations, put forward by Frank Jackson and Philip Pettit, which is designed to solve this metaphysical problem with social sciences in mind, fails to treat the problem in any importantly new way. Then I will propose a more modest approach, and show how it does not deserve the criticism directed at a prototype by Jackson and Pettit.

1 The problem
2 The pre-emptive model: an unsatisfactory proposal
3 The program model and how it collapses to the pre-emptive model
4 A more modest proposal

1 The problem
The business of social science is to explain the large-scale phenomena of society, and how they change over time. Examples of such phenomena are (1) declines in birth rates (frequently put down to rises in urbanization), (2) high crime rates (typically put down to high unemployment levels), and (3) widespread prejudicial attitudes amongst members of certain groups towards members of other groups (put down to more factors than can be listed in a parenthetical space). What species of explanation is appropriate for these phenomena? What class of facts should explanations of these phenomena draw on? There are bitter controversies over this matter between individualists and anti-individualists (sometimes called structuralists), which are routinely categorized as ‘methodological’ disputes. However, these disputes are manifestations of a general problem that all nonphysical sciences face: the problem of legitimating their existence alongside physics. And this problem concerns metaphysics, not methodology purely. This essay is concerned with the metaphysical subjects touched by this problem.

Individualists advance the proposition that there is a 'rock-bottom explanation of every large-scale social phenomenon which explains the phenomenon as due
solely to the beliefs and dispositions of actual or typical individuals and the situations to which they respond in accordance with their beliefs and dispositions. Here we are to understand the term ‘dispositions’ as ranging over agents’ reasons for action—the reasons they are disposed to offer in explanation of their actions. Individualism, as thus formulated, rests on the idea, expressed as early as the 1850s by J. S. Mill, that whenever a human being has genuinely taken action—rather than merely registered a reflex—the behaviour that results must be rationally explicable from a first-person perspective. And this idea rests on the metaphysical doctrine that a human agent taking action within a social context in which other human agents are also taking action, is not essentially different from one taking action in ‘splendid isolation’ (as it were). Thus:

The laws of the phenomena of society are, and can be, nothing but the laws of the actions and passions of human beings united together in the social state. Men, however in a state of society, are still men; their actions and passions are obedient to the laws of individual human nature. Men are not, when brought together, converted into another kind of substance, with different properties; as hydrogen and oxygen are different from water, or as hydrogen, oxygen, carbon, and azote, are different from nerves, muscles, and tendons. Human beings in society have no properties but those which are derived from, and may be resolved into, the laws of the nature of individual man. In social phenomena the Composition of Causes is the universal law (Mill [1872], Bk VI, Ch VII, Section 1, p. 573).²

Anti-individualists, of course, reply that the individualists’ constraint on admissible explanations is too restrictive. They offer a variety of reasons for rejecting the constraint, predicated on the metaphysical doctrines to which they subscribe. One conceivable reason is that, contra Mill, Composition of Causes vis-à-vis human aggregates does not hold universally. Another conceivable reason for rejecting the individualists’ constraint, related to the first, is famously put forward by Karl Marx.³ This is that the dependence relations between individuals on the one hand and social wholes on the other is much more complicated than would be suggested by the sort of analytic social science, à la physics, which Mill envisioned. The reason, according to Marx, is that—again contra Mill—social wholes are not posterior to individuals, either in time or in matters of influence. For

neither legal relations nor political forms could be comprehended whether by themselves or on the basis of a so-called general development of the human mind, but that on the contrary they originate in the material conditions of life ... In the social production of their existence, men inevitably enter into definite relations, which are independent of their will, namely relations of production appropriate to a given stage in the development of their material forces of production. The totality of these relations of production constitutes the economic structure of society, the real foundation, on which arises a legal and political consciousness. The mode of production of material life conditions the general process of social, political and intellectual life. It is not the consciousness of men that determines their existence, but their social existence that determines their consciousness (Marx [1970], pp. 20–21).

This dispute between individualists and their opponents is a special instance of the dispute between nonreductive materialists and reductivists, and this highlights the fact that the dispute is a manifestation of a general problem that all nonphysical sciences face: the problem of illuminating the grounds on which rests their independence of sciences presumed more ‘fundamental’—and so also the grounds of their unconditional right to exist alongside of more fundamental sciences. Nonreductive materialism, as a philosophical doctrine, has perhaps as many different forms as it has boosters, all embracing the following three propositions:

1. The world of experience is composed of ‘layers’, ‘levels’, or ‘strata’ of facts. The lowest, or most ‘basic’, of these is the microphysical, which is splendidly treated by microphysics. Microphysical facts ‘realize’ all facts at the higher levels.
2. Each (or at least one) nonbasic stratum is the preserve of an autonomous or independent discipline, which cannot be eliminated—not even in epistemological utopia—in favour of physics.
3. Facts and events at the microphysical level cannot be causally explained by facts or events at higher levels. In other words, there is no ‘downward causation’.

Propositions (2) and (3) are parallel protectionist doctrines: (2) declares the existence of provinces that must be protected from physics, ostensibly by the resources of the correct theory of explanation, while (3) declares that physics is protected from domains ‘above’ it, by a kind of causal closure to influences from ‘above’. The prerequisite of our being able to draw boundaries between domains—as is readily and routinely done by those who exercise administrative powers in university settings—is that we had better be prepared to protect these boundaries, within a comprehensive account of science as a whole.

Now propositions (2) and (3) are potentially in conflict, if accepted into a metaphysical framework in which causal explanation is synonymous with scientific explanation. For, as the reductivists will point out, if all facts are

---

¹ R. Miller [1978] argues persuasively that this formulation of individualism is both plausible and nontrivial, and attributes it to J. W. N. Watkins. This formulation is criticized, unconvincingly, in Ruben [1985], pp. 150ff. For a defence of individualism, see D. H. Mellor [1982].
² The individualists’ aversion to their opponents’ view sometimes takes the form of a charge that the opponents commit the so-called ‘fallacy of composition’. See R. Hardin [1982], esp. pp. 1–3.
³ Miller, op. cit., wisely argues that individualism must be formulated so as to disallow Marxists’ explanations from qualifying as ‘rock-bottom’ explanations, and illustrates how his formulation (which I am here adopting) does precisely that.
realized by microphysical facts (proposition 1), and if microphysical facts cannot be explained by facts at higher strata (proposition 3), then when respectable ‘ground-level’ explanations of all microphysical facts are put forward, how can there be space left over for an independent, higher-level explanation of higher-level facts or events? Nonreductivists might reply simply that it is not true there exists a ground-level explanation for each and every microphysical fact, so that the problem of space left over does not really arise. But surely nonreductivists do not wish to advance their nonreductivism on the strength of this proposition. Surely nonreductivists wish to make room for their higher-level enterprises, whether or not there is ‘explanatory space’ below.

We can illustrate this general problem, in a slightly different way, by exhibiting the central patterns of explanation in the very successful field of quantitative evolutionary biology. There one explains the prevalence of a particular trait within a population of organisms by drawing attention to a selection history of the ancestors of the present population, which features them as winning the competition with other organisms for reproductive resources. This is obviously not a ground-level explanation, as its friends will cheerfully admit. And, at their behest, we may well accept it as compatible with the true, ground-level, explanation of the phenomenon to be explained (the prevalence of a certain trait), which presumably involves the activities of subatomic particles and light rays, exclusively. But it is one thing to propose the compatibility of a selectionist explanation with the sort of explanation (presumed causal) that predominates in microphysics, and another thing altogether to declare their independence—particularly when one embraces the proposition that the higher-level facts, which a selection account aspires to explain, rest on facts at the ground level. And exactly the same thing holds true of the phenomena treated by cognitive psychology: it is one thing to propose the compatibility of a belief-desire style of explanation for some piece of behaviour with a microphysical or neurobiological one, and another thing again to assert their independence.

Returning now to social science, the reductivist or individualist may argue as follows. Since, as the nonreductivists themselves allege, all social facts are realized by facts about individuals, and since no facts about individuals can be causally explained by facts about populations, how, when respectable individual-level explanations of social facts are available, can there be space for an independent, higher-level explanation of those very same social facts?

2 The pre-emptive model: an unsatisfactory proposal

One way out of the difficulty faced by nonreductivists is, obviously, to reject proposition (3). The result of this surgery is what F. Jackson and P. Pettit call the ‘subversive model’ or ‘pre-emptive model’ of social explanation, which they trace to the likes of Marx and Engels—to certain discussions of false consciousness and ideology—and to Durkheim’s discussion of the causal potency of the social fact, as much as to the French structuralists, and Louis Althusser in particular, who proclaimed ‘the abolition of the human subject as the cost to be paid—in his case it appears to have been paid with some enthusiasm—for the insights of Marxist macrotheory, correctly understood.’

Jackson and Pettit are very critical of this model, writing of those who embrace it:

They must say that, contrary to our own intuitive sense of these things ..., by whatever instrumentality, a relatively abstract feature like increased urbanization, the social benefit of stratification, the class benefit of tolerance, or the ethos of capitalism must be capable of eliciting in individuals the psychological profile which guarantees that the social explanandum obtains (ibid., pp. 109–10).

Jackson and Pettit object to the subversive model for two reasons: (1) that macrofactors are very unlikely candidates for causes of their alleged effects; and (2) that the subversive model undermines folk psychology, because it suggests that macrofactors of a social nature have direct and unrecognized causal influences on our psychological make-up, at least in so far as the subversive model posits the existence of ‘unfamiliar limits on our sensitivity to evidence and the like; it will subvert, in whatever measure, the manifest image of the human being as an intentional system’ (pp. 110–11). I mention these criticisms because, as I shall argue, Jackson and Pettit’s own model collapses upon further analysis into the subversive model, and so is subject to these two criticisms as well.

3 The program model and how it collapses to the pre-emptive model

Friends of the subversive model do away with proposition (3). But what can they say who either wish to, or must, for whatever reasons, embrace all of (1)–(3) in full strength, but wish nevertheless to reserve the prerogative to affirm the notion that there is some independent scientific value to structural explanations? Jackson and Pettit claim to have the solution for these unfortunate friends of the social sciences, and claim that this solution generalizes to solve as well, and on a very abstract level, the overarching problem of this

---

4 Jackson and Pettit [1992b]. Unless otherwise indicated, all references to Jackson and Pettit will be to this article.

5 ibid., p. 111.
essay.\(^6\) They call their solution the ‘program model’. And it has already been cited in support of certain philosophical positions about the appropriateness of higher-level explanation in psychology.\(^7\) According to Jackson and Pettit, this model is compatible with all of (1)–(3), but still affirms the proposition that structural explanations nonetheless provide information not found in individualistic explanations. And among its alleged virtues Jackson and Pettit claim it can also answer the ‘missing-mechanism’ criticism of functional explanation, formulated by J. Elster.\(^8\) As I shall argue, however, the program model is just the subversive model in a new guise.

To appreciate the program model we must first distinguish between a causal process, and what Jackson and Pettit call a program. Jackson and Pettit do this by distinguishing between process explanation and program explanation. A process explanation functions by bringing forward information about a causal chain at work in the actual world, and a program explanation functions by bringing forward information about causal chains at work in different possible worlds. They illustrate the distinction by the example of a closed flask that breaks as the temperature of the water inside it is brought very slowly up to and maintained at boiling point; the flask breaks as the vapour pressure increases. They write that the ‘salient microconsideration is that a certain molecule or group of molecules collides with a molecular bond in the surface of the flask at a sufficient velocity to break it’ (p. 117). But—as they formulate the problem to be solved—if this is the cause, how can it be interesting from the point of view of explanation to bring to attention macro facts to do with temperature, such as that it is at boiling point? Their reason is to assure us that temperature explains the flask’s breaking ‘simply because the rise in temperature means nothing more or less than that the rate of motion of the water molecules will increase, and if the rate of motion increases then it is more than likely that some molecule will have the effect explained’ (p. 118).

The temperature explanation of the flask’s breaking is a program explanation, because it appeals to a property (temperature) which ensures (makes it suitably probable) that there will be a causal process at the lower level (namely, a molecule with sufficient velocity) which results in the fact to be explained, but this programming property does not itself cause that fact. The factor that programs for an effect does not get to produce it, at least not in the sense in which the so-called producer state does so. It non-causally arranges things (it means that things are arranged) so that there will be such a producer state—maybe this, maybe that—available to do the work.\(^9\) Programming is causally relevant, but not causally efficacious, since it does not itself figure in a causal process.

The distinction between program and process, according to Jackson and Pettit, forces us to acknowledge two categories of factors, one embracing or including the other: the causally efficacious, on the one hand, and the causally relevant, on the other—with the latter embracing the former. The ‘merely’ causally relevant will then comprise factors which fall in the second category but not in the first. Which of these categories a given factor falls into—whether into the efficacious class or the merely causally relevant class—is a relative matter: it depends on which stratum of the material world we are currently dealing with. And so long as we are able to draw a boundary between these two types of factors, relative to a given level, we can (at any given level) acknowledge two kinds of interesting scientific explanations.

Applying this idea to social phenomena, the program model proposes to explain as follows: that certain structural factors, such as for example increase in unemployment, ‘program for’ the appearance of certain social facts (criminal incidents, not through bringing them about causally, but through ensuring that some or other causal factors sufficient to produce the effect (elevated crime rates) will appear. The structural factor explains, relative to the level at which we are treating the phenomenon, not by being efficacious but by being merely relevant.

The program model achieves its appeal by suggestively drawing on a computer analogy: the distinction between a set of coded instructions and the real-time computations which result from their implementation. There is indeed a genuine distinction between the code and the sequence of computations which result when the code is implemented—one non-negligible difference is that the first must precede the second in any narrative of the relevant sequence of events. However, not even this gentle differentium is available to distinguish program from process in the cases Jackson and Pettit call to attention. Temperature does not—or at any rate need not, according to Jackson and Pettit—appear as an antecedent in an efficacious causal chain, nor need it contribute to the efficaciousness of other factors in such a chain. As I shall now argue, the distinction Jackson and Pettit would like to draw must remain in the ordinary-language domain: it cannot be given safe passage into the precise language of scientific philosophy where the notion of causal process is treated with great (and sometimes also very technical) sophistication. There the distinction between program and process disappears, like the emperor’s new clothes. And with the disappearance of the distinction goes the ability to avoid downward causation.

Suppose, on the one hand, we take the counterfactual account of causation—which Jackson and Pettit would appear to favour (p. 119). How shall the causal relevance condition be stated? And how, subsequently, shall the boundary between the efficacious and the merely relevant be drawn? Jackson and Pettit wish to say that certain structural factors raise the chances of

---

\(^6\) Their solution is applied to the general problem in Jackson and Pettit [1990].

\(^7\) R. Wilson [1994].

\(^8\) Elster [1979]. Pettit’s reply to this criticism is found in his [1996].

\(^9\) P. Pettit [1993], p. 37.
certain lower-level events. One natural way of formulating this idea within a precise counterfactual framework for explicating causation, would be to say that the explanandum would not (or most likely would not) have occurred if the ‘programming’ factor had not been present. However, this statement automatically qualifies the programming as a causally efficacious factor, since causal efficacy is defined simply as counterfactual dependence within a counterfactual account of causation. Thus, unfortunately, the program model simply collapses to the subversive model, which Jackson and Pettit are at such pains to find faulty.

(And if, more explicitly still, we take a closest-worlds account of counterfactual matters—as, for example, David Lewis does, who is cited with apparent approval by Jackson and Pettit at p. 119—we can show that the distinction between efficacious and merely relevant causal factor disappears immediately. For on the closest-worlds account, counterfactual dependence, for which all causal relations—and the efficacious ones in particular—go proxy, is a way of representing facts about different possible worlds—namely, those sufficiently like the actual one in the factors being examined. Thus if Jackson and Pettit are serious about the distinction between process and program being founded on the distinction between possible and actual chains of events, this distinction must rest on whether trans-world considerations are being brought forward. But since, according to Lewis’s view anyway, the distinction between causal chain of events and noncausal chain of events must itself be founded on trans-world comparisons, there will be nothing that qualifies as a causally relevant factor which will not also qualify as a causally efficacious factor—whether we speak in an absolute sense of that term or in the relativized sense Jackson and Pettit favour.)

Perhaps Jackson and Pettit will suggest that we ought not to convert the statement about structural factors raising chances into a statement about counterfactual dependence, and simply view the raising of chances as independent facts about the world, in a class of their own. This will avoid the problem just raised, but will raise another. If the statement, that a certain structural factor raises the chances of a certain lower-level factor, provides independent causal information from any statement to the effect that a certain nonstructural factor serves as an efficacious cause, then it will be false to say, as counterfactualists prize the ability to do, that an efficacious cause raises the chances of the explanandum. For if the raising of chances is independent of the bringing about, then bringing about cannot be done through the raising of chances, any more than the raising of chances can be brought about through an efficacious cause. Consequently, raised chances cannot serve as evidence for the efficaciousness of any factor.

But suppose, on the other hand, that Jackson and Pettit prefer a probabilistic account of causation, in which a factor is defined as a positive (efficacious) contributor towards production of a given explanandum if (among other things) it raises the probability of the explanandum. But the prospects here for a categorial distinction between program and process are now even worse, since probabilistic relations are here definitional of efficaciousness, and Jackson and Pettit wish to define mere causal relevance as well in terms of raised probabilities. So since, as Jackson and Pettit will have it, a programming factor elevates the probability of the effect, it automatically qualifies as efficacious too, if we adopt the probabilistic account of efficaciousness. On such an account, the difference between being merely relevant and being efficacious would be, at best, a difference in degree, not a difference in kind.

Will it help to take an INUS account of causality, à la Mackie, which has a place for probabilities in it, but not in the definition of (efficacious) cause? On an INUS account, a factor will be causally efficacious if and only if it is one of a group of conditions which are sufficient for the explanandum (thus a cause is an Insufficient but Necessary part of a condition which is itself Unnecessary but Sufficient for the alleged effect). How shall we, on Jackson and Pettit’s behalf, define the quality of being merely relevant? Jackson and Pettit will have it that a relevant factor raises the chances of the explanandum. This can happen, on an INUS account, only by way of the factor in question being one of a group of conditions sufficient for the explanandum. Thus once again we find no means of drawing a categorial distinction between the efficacious and the merely relevant: whenever a factor qualifies for being relevant, it qualifies automatically for being efficacious as well. Jackson and Pettit may well reply here that a merely relevant factor simply raises the chances of an efficacious factor appearing, without actually lying within a group of conditions sufficient for the explanandum. This seems promising. But the problem here is that the raising of chances, within an INUS account of causality, rests on the distribution of efficacious factors, and is not independent of efficacious factors. Just as in the counterfactual account, the handling of raised chances within an INUS account, as a function of efficacious factors, is predicated on the decision to treat raised chances as evidence for efficaciousness. And this, once again, is contrary to what the program model requires. The program model requires that the appearance of an efficacious factor be independent of the appearance of a merely relevant one—otherwise drawing attention to the merely relevant factor would add no information, becoming redundant, once the efficacious factor has been revealed. But like the counterfactual account, the INUS account has the dependence relation going in the other direction: efficacious factors

10 Obviously I am oversimplifying the probabilistic account of causation. However, a more nuanced probabilistic account will not help Jackson and Pettit here, since their formulation cannot exploit nuances to escape its troubles.

11 J. L. Mackie [1974].

12 For a good discussion of the differences between INUS products and probabilistic accounts, see D. Papineau [1985], pp. 57–74.
determine the chances of everything else, and are not themselves affected in their probabilities of occurring by any other type of factor. And this is contrary to what the program model requires.

Finally, Jackson and Pettit might join forces with those who oppose probabilistic conceptions of causal processes out of antireductivist principles, on grounds that efficacious processes are somehow more fundamental even than objective chances. Thus Jackson and Pettit might elect to take the notion of efficacious causal process as primitive and irreducible to probabilistic relations, and define the notion of relevance (as they have done) in terms of raised chances. This is most promising for their enterprise. (Jackson and Pettit will be alone, however, in requiring the facts of probability to be independent of the facts of causation, as the antireductivists still wish to say that the raising of chances is done through efficacious factors.) If they do this, however, the judgement that some factor is causally efficacious, and not merely relevant, will always be unchallengeable. It does not give the opponents a sporting chance at dislodging the program model from a place of honour. If Jackson and Pettit propose to distinguish between the program and the process in this way—by fiat—this is a solution to their problem. But it has all the advantages of theft over honest toil.

In a footnote that proposes to dispel these difficulties (p. 118), Jackson and Pettit remind the reader that the factor invoked in a program explanation may be ‘individuated so that it is the same event as the cause, The important point then is that the property highlighted is not causally efficacious in the same way as the property invoked in the process account’ (emphasis mine). But the problem to which I am here drawing attention is that of how to explicate this phrase ‘not in the same way’, in such a way that the distinction which emerges can be used to draw a precise boundary between the two categories of factors they wish to distinguish, even if the boundary is what it is only in relation to the ‘stratum’ containing the fact we need explained. As they themselves state, their position requires that there be ‘two grossly different sorts of information that can be provided’, within a causal account, about a series of events. They wish to call the first type of information contrastive, because it helps us differentiate between possible worlds, whereas the second sort of information (which structural explanations provide) they wish to call comparative, because it helps us to discern constancies across possible worlds. So that the explanation of the flask’s breaking in terms of this or that molecule colliding with the surface gives us contrastive information, whereas explanation of the flask’s breaking in terms of the temperature gives us comparative information. While Jackson and Pettit have provided the labels for the categories they wish to identify, they have not provided the terms by which we can decide how to apply these labels in particular cases. And, as I’ve just argued, no account of causation currently on the market will assist them in doing this. Whichever currently marketable position we embrace vis-à-vis causation, we shall be able to make no categorial distinction between program and process, in the way Jackson and Pettit would like, even if we wish to do this in a level-relative way. The reason is that the respectable positions on causation all treat causal factors as contributors in degree, rather than in kind, to production of the effect, and if not a contributor in degree, then no causal contributor at all. There is currently no respectable account of causation that separates factors by category. Thus there will be nothing which qualifies as a causally relevant factor, on any view of causality currently on the market, unless it also qualifies as (part of) a causally efficacious one.

My point, then, is that there is a very strong thread woven into each of the reputable accounts of causation currently on the market—a thread which is not likely to be left out even if an antireductionist account of causal processes is accepted. This is that the raising of chances is done through a distribution of efficacious factors—that, in other words, the facts of probability rest on the facts of efficacious causation, and as a result, are not independent of them. Now it is precisely this thread which Jackson and Pettit would like to draw out of the account of causation. Thus their account is incompatible with any of the accounts of causation currently on the market, or any successor likely to come along any time soon.

It may help to underscore this point if I show how Jackson and Pettit’s own examples do not serve their need to illustrate, without giving us a new account of causation, how the condition of relevance can be met without the condition of efficacy also being met. The closed-flask example is their most faithful. The temperature factor in the structural explanation of the flask’s breaking is supposed to satisfy the relevance condition, but not the efficaciousness condition, by satisfying all of the following: (i) it is a factor which is efficacious only if another (more basic, microscopic) factor is efficacious; (ii) it does not appear in a chain comprising efficacious microfactors only; and (iii) it does not contribute to the efficaciousness of a link in such a chain. However, satisfaction of (i)—(iii) does not stand in the way of the temperature factor itself qualifying as an efficacious factor, contrary to Jackson and Pettit’s belief. For it does not rule out the possibility of a chain of efficacious causal factors as follows. (I) microfactor A ⇒ macrofactor M ⇒ microfactor B. (Please read ‘⇒’ as signifying efficacious causation.) Macrofactor M, in this chain, might also satisfy all of (i)—(iii), simply by logically entailing the existence of the following chain: (II) microfactor A ⇒ microfactor m ⇒ microfactor B. Thus condition (i)—to the effect that the temperature is efficacious only if another, more basic factor is efficacious—is satisfied, along with (ii) and (iii), by
macrofactor M and microfactor m both being efficacious. The existence of chain (I) is, of course, what Jackson and Pettit wish to rule out, as they oppose the pre-emptive model, but nothing in the program model prohibits it. (Condition (ii) is perhaps Jackson and Pettit's best reason for thinking that the temperature factor in their example does not qualify as efficacious. They believe—correctly, and uncontroversially—that macrofactor M cannot appear in a chain of microfactors only. Logically speaking, however, this condition can be satisfied at the same time as chain (II) is satisfied, and at the same time as the third condition is satisfied—particularly if, as Jackson and Pettit say, 'what is invoked may be individuated so that it is the same event as the cause.' ) Thus Jackson and Pettit have failed to carve out space for a set of conditions which fall into the category of relevant, but not into the category of efficacious. And having failed at this, they cannot achieve the aim of avoiding a pre-emptive account of social explanation, for they have failed to show how chain (II) is incompatible with chain (I).

Perhaps it will be said that, just as the program model does not require the existence of a metaphysical ground level or 'bottom' to the universe, and just as it does not require the distinction between efficacious and merely relevant to be absolute rather than level-relative, so it also does not require that there be fine lines drawn, for each stratum, between the efficacious and the merely relevant. Efficaciousness shading into mere relevance is acceptable, and will catch the important facts about explanation to which Jackson and Pettit draw attention. For very few boundaries can be drawn easily in a sharp fashion, but this doesn't matter so long as we have clear cases on either side. I reply that it is a mistake to let Jackson and Pettit off the hook on grounds that efficaciousness might come in degrees, shading into mere relevance. That it might do so is neither here nor there for our purposes. For our purposes we need to notice that—as Jackson and Pettit themselves realize—the program model presupposes (in Jackson and Pettit's own words) 'two grossly different sorts of information that can be provided' at each stratum of the world one may care to treat, whether there is a bottom to the universe or no. Shades between efficacious, on the one side, and merely relevant, on the other, will not do. And the reason is that the information presented at a higher level has to be independent of the information presented at the lower level. One can't achieve the independence of social science from physics—indeed, independence being an all-or-nothing affair, not a thing of degree—if efficacious shades into merely relevant by degrees.

Of course it is open to Jackson and Pettit to reply by saying that the distinction between program and process should not disappear in a respectable treatment of causal explanation, and that therefore it is to the embarrassment of philosophers of causation that the distinction has no place in their treatments. For Jackson and Pettit's program model has surely caught an important relationship between two types of explanations of the same event—a relationship for which (if I am right that accounts of causation currently on the market do not discriminate types of causal contributors by category) philosophers of causation have no suitable account. If this is the reply—that Jackson and Pettit's program model exposes a weakness in all the accounts of causation currently on the market, and reveals that we need a radically different type of treatment of causation—then the program proposal is much more sweeping than its unprepossessing presentation would have suggested. Even so, this reply concedes that the program model is still in the promissory stage, awaiting a more satisfactory account of causation—which is, in slightly different terms, the complaint I am lodging.

4 A more modest proposal

I propose more modest means of disarming the social scientist's problem, created by propositions (1)–(3) (stated at the beginning of this essay): a certain type of selectionist proposal for construing the structural explanations produced by social scientists. Jackson and Pettit do their best to retain all of (1)–(3). The pre-emptive model does away with (3). My suggestion is to do away with (1), or at least to dilute it, in favour of the proposal that there is only one stratum of reality, consisting of interdependent facts—or, at least, that there is no boundary between the individual level and the social level. In other words, the domain of empirical science cannot be divided into islands of independent realities. Such divisions would have to be made on the principle that the proposed strata admit of grossly different, and incombines, types of explanation; yet these gross differences and incombines are fewer in number than might be supposed, if indeed there are any. In particular there are no gross differences between what is appropriate to explaining individual behaviour and what is appropriate to explaining phenomena on the social scale. Thus my proposal forces us to acknowledge that administrative realities, of the sort that provide the provost of a university with good reasons for dividing the world of facts in convenient ways, do not amount to metaphysical realities. With (1) thus diluted or done away with, the combination of (2) and (3) ceases to be so troublesome. For fewer border protections will be necessary if there are fewer borders to protect. Here is how to implement this proposal.

The selection pressures spoken of in biological sciences are typically due to natural conditions—either to scarcity of resources for reproduction or to the need to survive encounters with predators. Natural conditions exert natural selection pressures. Natural selection operates by selecting those

---

16 I owe this point to the very helpful comments of an anonymous referee.
characteristics which enhance an individual’s chances of producing descendants into future generations. Now according to my proposal there is, in addition to natural selection, such a thing also as social or cultural selection, and that the pressures which give rise to this type of selection are exerted in social, as contrasted with natural, settings. Of course this idea is not new; it is the legacy of those social scientists who, in the last few decades, have been developing a scientific theory of cultural change, in the spirit of evolutionary biology. These pioneers are aspiring to contribute to quantitative and mathematical anthropology and social science. I wish to put their insights to work in elucidating a means of disarming the overarching problem of this essay.

Cultural transmission of a characteristic is typically brought about through social interactions, which operate to exert a selection pressure. Sometimes this selection pressure involves competition for something—for example, an ideal of life or lifestyle—the attainment of which may actually reduce an individual’s chances of producing descendants into future generations. Thus natural selection and cultural selection may operate in opposite directions, one opposing a certain trait while the other favours it. One example of the operation of cultural selection, made famous by L. Cavalli-Sforza and M. Feldman, is the evolution of the disposition to have fewer offspring. As Cavalli-Sforza and Feldman explain, this trait is transmitted through a system of social reinforcements, applied on an individual-to-individual basis in the form of approval or disapproval. But it admits of both horizontal and oblique transmissions (transmissions across genetic or family lines), as well as vertical transmissions (transmissions from parent to child). Cavalli-Sforza and Feldman show that such a selection process can explain the dramatic decline in birth rates in nineteenth-century Western societies, in which average family size went from five to just over two, while strictly biological models of fitness can never explain this phenomenon (under the circumstances prevailing in nineteenth-century Europe, for instance), because the trait of rearing only two offspring reduces biological fitness under the prevailing conditions, while the trait of rearing five raises it. (The prevailing conditions include a drop in death rates due to epidemics and famines, as well as overall decline in mortality rates.) Thus the idea of cultural evolution can help to explain, where natural selection cannot, why characteristics that don’t contribute to more offspring in future generations can, nonetheless, become prevalent.

How does the selectionist proposal I am hereby urging stand vis-à-vis propositions (1)–(3)? The selectionist model addresses the individualist worry that large-scale behaviour cannot be explained in terms of many prudent individual responses to prevailing conditions, by bringing to attention selection pressures of a previously neglected kind, which help to make certain courses of action understandable (prudent) at an individual level. Even so I have not yet said how my selectionist model answers the very general problems facing nonreductive materialism—for doesn’t my proposal ultimately draw attention to evolutionary biology’s own difficulties in legitimating its unconditional existence alongside physics? It does, but it also draws attention away from the difficulties of social science, as such. And this is the point. What the selectionist proposal does, consequently, is to suggest a rather nice procedure for disarming the general problem we are discussing. But it does not guarantee that this procedure can be repeated with universal success.

The selectionist construal of structural explanations commends the idea that there is no dividing line between the individual and social levels, since evolutionary principles apply to both. In other words, it recommends that the material world be divided into strata only according to whether there are limitations to the application of patterns of explanation. On the basis of that principle, the selectionist model finds no grounds for making certain proposed divisions, including the division between the individual and the social. And this minimizes the difficulties of embracing proposition (3): the spectre of downward causation does not loom where we do not acknowledge a hierarchy of strata. Thus the selectionist construal suggests that we can reject entirely the proposal that there are real divisions in nature as discussed in proposition (1), the problem we have been illuminating will be completely and satisfyingly dissolved. But it does not guarantee success. And this is as it should be, for the proposition that there are no independent domains within the empirical sciences is itself an empirical proposition.

Jackson and Pettit believe that the selectionist construal of structural explanations takes a favourable view of proposition (3): it ‘presents a picture under which the structural factors may be less basic causes and still serve to provide autonomous and interesting explanations’ (p. 112). Jackson and Pettit would interpret, incorrectly, the theory put forward by Cavalli-Sforza and Feldman as stating that a selectionist pressure (in this case, a certain widespread practice of approval and disapproval among one’s family and peers) is perhaps a ‘cause’, though a less ‘basic’ one, of the effect to be explained (the average family size shrinking from five offspring to just over two). The reason for their mistake is, as I see it, that they believe, incorrectly, that theories which invoke social selection are invoking group selection. Thus they write:

18 Ibid., pp. 180–9.
19 There is an argument in Pettit [1993] (pp. 159f.) in support of the proposition that only a model with group selection has the right sort of credentials, but this argument is flawed, as the Cavalli-Sforza–Feldman account is a counterexample.
were bound to fall or stand, disappear or survive together; (2) these differed in certain aggregate features, in virtue of their members differing, whether for genetic or other reasons, in behavioural dispositions and patterns; and (3) over the period in question there must have been a relatively unchanging environment such that a society’s performance in the survival stakes would have consistently been favoured or jeopardized by the same sorts of aggregate features (p. 114).

The claim that a selectionist model for interpreting structural social explanations must presuppose group selection is incorrect, as the example exhibited here shows. For Cavalli-Sforza and Feldman’s theory is a social-selection theory only in so far as it posits the existence of trait-transmission mechanisms that are social in character. In other words, Cavalli-Sforza and Feldman explain decline in family size by exhibiting a certain social force, in addition to all the natural forces, against which individuals must navigate. According to their explanation, an individual seeks (among other things) to minimize the balance of social expressions of disapproval over approval. Thus theirs is a theory of social selection not because it claims that the units of selection are social wholes, but because it posits pressures exerted through social interactions, as well as a system of trait-transmission that is social in character. Thus the selectionist construal of structural explanation does not suffer the deficiencies documented by Jackson and Pettit.

Acknowledgements
I am indebted to Barry Smith, Julian Barbour, Anna Maidens, the audience in attendance at the British Society for the Philosophy of Science, Merton College, Oxford, September 1997, and an anonymous referee for this journal for helpful discussions, comments and suggestions on earlier drafts.

Philosophy Department
State University of New York at Buffalo
Buffalo, NY 14260
USA

References


Irzik, G. [1996]: ‘Can We Reduce Causes to Correlations?’, British Journal for the Philosophy of Science, 47, pp. 249–70.


Wilson, R. [1994]: Causal Depth, Theoretical Appropriateness, and Individualism in Psychology, Philosophy of Science, 61, pp. 55–75.