Formal Sociological Theory

Possibilities and Challenges

R.J. Anderson
Horizon Digital Economy
University of Nottingham

W.W. Sharrock
Department of Sociology
University of Manchester

Over the past few years, we have returned several times to the suggestion sociology should be constituted on models derived from the natural sciences. This is yet another essay on the same theme. Though the various pieces are related in many ways, they were written to be self-standing. At some point we might have the motivation and energy to integrate them into a single coherent argument but, as with St. Augustine and chastity, not just yet.

We would like to thank an anonymous reader for a number of helpful comments on the previous draft.

Version Final

I. INTRODUCTION

We doubt William Thomson thought about sociology very much. And if he did, we doubt he thought very much of it. Yet Lord Kelvin (as he later became) has a lot to answer for. His imperious aphorism dismissing the value of non-quantified knowledge has been regularly used to encourage sociology to adopt methods and procedures modelled in particular on Thomson's own discipline of physics,[1] the intention being, of course, for sociology to acquire the kind of knowledge which would allow it also to pass muster as science. Although the debate over 'sociology as science' has largely died down of late, it hasn't entirely gone away.[2] For example, a recent ESRC Review Panel[3] advocated adopting "the Kelvin wedge" of quantification to remedy the deficiencies (as it saw them) of contemporary UK sociology. The trouble is the proponents of a fully scientific sociology have rarely, if ever, bothered to consider whether they could actually satisfy the conditions which physics has imposed on itself and which make it the discipline it is nor, if they could, whether they would want to. The point of quantification, we presume, is to improve the sociology we have, not to give us a new sociology which nobody would want to do.[4]

What, then, are these conditions which physics has imposed upon itself? We think they fall into four broad buckets.

1. A presumption that the apparent heterogeneity of physical phenomena is underpinned by a discoverable order. There is surface chaos and underlying pattern.
2. A presumption that this pattern is only discoverable through a phenomenological reduction within which empirical phenomena are treated as systems of objects defined by measurable properties. These systems exhibit conformity to deterministic, reversible laws and are organised as complexes of causal mechanisms and their effects.[5]
3. A requirement that descriptions or models of systems of objects should be as complete and as precisely defined as possible. Rigorous identification of the components and

---

1 An alternative choice is proposed by LEIBERSON AND LYNN (2002)

2 Jerald Hage's edited collection (HAGE 1994) seems more like a premature wake than anything else.

3 See ESRC (2010).

4 We appreciate that these days no-one has great expectations regarding proposals to formalise sociology. However, amongst the minority who do see formalisation as desirable, it is supposed that this would introduce tangible improvement. For those thinking thus, it is that not enough sociologists are trying to turn their discipline into a science which is the major obstacle to progress. We are not discussing formalisation because we think remains a major force in sociology, only because it represents a position that persists.

5 We recognise this Laplacean characterisation was disputed throughout the period in which Classical Mechanics dominated physics and does not apply to Quantum Physics. That has not stopped it being widely promoted as the model of science sociology should aspire to. PETER SCHOENEMANN (1994) has an incisive discussion of the "reasonableness" of social science's failure to rise to the challenges of measurement compared to the "unreasonableness" of the effectiveness of mathematics in the natural sciences.

relationships being modelled demands equally precise and rigorous measurement systems.

4. A requirement that standardised protocols should be followed to render general descriptive propositions in functional form. Such formalised statements should exhibit securely grounded estimations of variables and their parameters.

Although the majority of sociology probably accepts the first presumption as its own *modus operandi*, it is the conditions that follow which make life awkward. First of all, no matter what their methodological commitments might be, all forms of sociology want to find space for explanations incorporating the interpretive character of social action. Laplacean determinism sits ill with this. Second, while sociology has problems enough with what might be thought of as 'feed-forward' generalisations about the effects of causes (think of the struggles we have defining the consequences of economic development, social deprivation, differential socialisation practices and so on), any attempt at 'feed-back' explanation of the causes of effects is fraught and fragile (SMITH 2014). Attempts to trace the origins of modernity have foundered as have attempts to explain political quiescence in the face of extensive and accelerating social inequality, and so on. Explanations that barely work for the effects of causes fail so miserably on the causes of effects that the resulting explanatory gap can only be closed by ideological fiat. Third, our inability to specify our models at a level of resolution (Einstein's famous "as simple as possible, but no more") which identifies and traces sufficient of the mechanisms at work to provide an end to end causal explanation is widely acknowledged. At various critical junctures we are forced back on accounts based in the possibility of 'action at a distance', something which is anathema to Classical Mechanics and the adoption of which by some versions of Quantum Theory led to Einstein's scathing rejection of them as "spookey". Finally, because so much sociology is articulated in non-formal, non-quantitative terms, formalisation itself has been a continuing challenge. Quantification requires interpretation within a mathematical formalism of the number system to provide formally secure measures of the properties we are interested in.[6] If the translation of sociological data into such formalisms is not possible without processing out the properties of descriptive richness and contextual detail we prize in our accounts, wouldn't we be forced to ask if they are worth producing?

In reality, the demand for quantification is a demand for a physics-like methodology. In previous discussions,[7] we have made much of sociology's 'missing theory of measurement' as a constraint on its ability to satisfy the requirements for being a science. However, we need to remember formalisation does not necessarily mean quantification. In this discussion we set aside matters of measurement to focus exclusively on the question of formalisation itself and the construction cogent, formal (but non-quantitative) propositions from non-formal, discursive sociological descriptions. Our argument is not

[6] It is important to recognise the usual sociological practice of counting frequencies is no more formal than the farmer's counting sheep. Unless these frequencies can be shown to be measures of some property we are interested in, they are 'quasi-formal' at best.

[7] See ANDERSON AND SHARROCK (2013, 2014, 2015)

that formal theories in sociology are impossible. Rather, we suggest that as currently envisaged a move to formally specified theory must involve solving a number of difficult and unresolved (perhaps unresolvable) problems as well a number of trade-offs we might be reluctant to adopt. Certainly, the end result of overcoming the problems and adopting the trade-offs would undoubtedly be a very different sociology to the one we currently have. Since sociology is as it is for very good reasons, we need equally good reasons to want to change. At the moment, no such reasons are on offer from the proponents of formalised theory.

To demonstrate the basis of this claim, we adopt the following strategy:

1. First, we try to pin down what is meant by formalisation and the benefits it might offer sociology. We will see that one of the major hoped-for outcomes is a reduction in sociology's disorderliness consequent on a regimentation of its forms of expression, definition of topics and specification of problems.
2. Second, we examine the challenge of translating any non-formal into formal theory. Not very originally, we call this the 'indeterminacy of translation'.[8] This problem is important because if a move to formal theory is to make sense, one of the requirements, presumably, would be the retention of 200 years of extant sociological knowledge. Having found a formal language for sociology, we would want to translate classical and contemporary theory into it so that we could build on what we have. Without such translatability, a move to formalisation will mean more or less starting all over again. Not a prospect many will relish!
3. We then consider a recent set of proposals which highlight the requirements which formal natural science theories impose on themselves when constituting causal explanations. This is important because one of the main reasons for wanting formal theory is to be able to offer causal explanations which will stand comparison with those of the natural sciences. The lesson we will draw from this part of our discussion is that currently sociology cannot satisfy the conditions as set out and moreover, given its framing, it is genetically incapable of adopting the assumptions necessary for it to do so.
4. Finally, we turn to an example of formal theory building. We review the theory's presuppositions, theorems and findings and suggest they fall short of the required generative adequacy which is the test of axiomatic theorising. In addition, as currently constituted, this theory offers no clear pathway to the scope of explanation which the sociology it aims to encapsulate actually addresses.[9]

---

8 We have borrowed this term from the work of the philosopher W.V.O Quine where it is used to raise issues for formal theories of meaning. Clearly, ours is a very different usage of it. See QUINE (1969).

9 Note we are not saying that a formal theory's presuppositions have to be 'realistic' in some sense. Since they are idealisations, how can they be? What we are saying is that once elaborated, the theory should give us back the social phenomena it analyses in descriptions we recognise to be apposite and convincing (or even true!).

Overall, the position we come to is that despite the energy they have already expended, those advocating the formalisation of sociology still need to provide detailed demonstrations of how the challenges we identify (and others of similar ilk) can be overcome and of how the sociology which results would be an order of magnitude enhancement on what we have now. The rhetoric of the debate on formalisation has largely been exhortation and admonition. We would rather be offered specification and demonstration. The reason for choosing the examples we do is that at least they provide that.

II. FORMALISATION

In this section, we look at the structure of formal arguments and some of the methods for building formal models. For convenience, we use quantitative examples. However, as we have said, quantification is a type of formalisation not definitive of it.

FORMAL ARGUMENTS

As a working definition, we will say that theories are collections of related propositions.[10] The relationship among these propositions is simply that some are stipulations or postulates while others are valid inferences from those stipulations. The rules for making such inferences are those of formal deductive logic.[11] Deductive arguments can and often are carried on using natural or ordinary language. However, because ordinary language is put to multiple uses apart from the enunciation of propositions, the ambiguity of many key terms (think of the many different ways a common term like 'table' is used) means it has been thought preferable to translate propositions into a *fully defined* and hence unambiguous notation. This translation is what is meant by 'formalisation'. The purpose of formalisation is to lay out the elements and steps in a propositional structure (that is, its argument) so as to secure and assess its deductive integrity, its transparency and its surveyability. A *rigorous* theory is all of these. The notation used may be any arbitrary string of marks providing the rules for their use is defined. While anyone could invent their own notation and lay out theories using it, that would be a wildly inefficient and inconvenient way of carrying on disciplinary discourse. Instead, depending on the purpose to which the formalisation is to be put, it has become usual to adopt the notations of either one or other branch of formal logic or one or other branch of mathematics. The point of applying this

---

10 For early discussions of formalisation in sociology see FURFEY (1954) and FREESE (1980). Although now ancient, HOCHBERG (1958) is still a superb summary of the requirements of scientific formalisation. HEYLIGHEN (1999) sets out the advantages and disadvantages of formalisation.

11 This discussion is not designed to be an introduction to deductive reasoning. We will assume an understanding of the rationale for Formal Logic though not a familiarity with any of its technicalities. In any case, we will not stray beyond an elementary use of first order logic. The examples we give are simple because we want to bring out what is going on in them not make claims about the phenomena themselves. With clarity about the structures involved, we can turn to the complexities of actual sociological theories.

notation is to shift focus away from the content of the theory (that is, the empirical referents of terms for the objects and relationships which the theory discusses) to reveal the deductive structure, *the form*, of the argument. Formalisation is in service of the specification, explication and validity of deductive reasoning not discoveries about empirical phenomena.

By way of simple illustration, we will use an example any parent (or Grandparent!) will readily recognise. Nell and Job are having strawberries. After two rounds of spooning fruit from the bowl into their dishes, the usual squabble about the fairness of portions starts. Nell's dish contains nine strawberries and Job's contains seven. Carefully we add six strawberries to Nell's dish and eight to Job's (and eat the odd one ourselves). The children have the same number of strawberries. We could represent what we have done by using numerals as a notation for the strawberries we have distributed.

$$9 + 6 = 7 + 8 = 15$$

So far so trivial. Notice, though, that our use of numerals is as *tokens* for spoonfulls of strawberries. We have said nothing about the strawberries themselves or the properties of these numerals. If we want to talk about 4, 6, 7, 8, 9 as numbers, we have to stipulate this by saying something like "Let 6, 7, 8, 9, be numbers" and rely on our informal understanding of what numbers are. As will become important later, this means all formalisation depends, at least to some extent, on undefined terms.[12]

What would be involved in giving a formal description of our use of numbers in distributing strawberries? That is, how might we describe the general process of adding any set of marks standing for quantities and not simply the numerals we have used? The standard approach is to use an algebraic notation such as:

$$a + b = c$$

We can start by saying "Let ***J*** be the set of natural numbers of which a, b, c are members", where the signs 'a', 'b', 'c' do not refer to any particular numbers. Strict formalisation would require us to define, 'set' and 'member', develop some associated notation and, no doubt, a lot of complex logic. For the moment we will leave these terms informally defined. The defined notation is, then, an *algebra for arithmetic*. This algebra is the formalism we are using to represent the form of our counting.

With this algebra, we can start to set out some assumptions about how numbers work; or to use the jargon, specify some axioms. If we have the numbers a, b we can stipulate:

$$a + b = b + a$$

This is known as **the axiom of commutation.**

---

[12] We could, of course, try to give a formal definition of number. However, this too would likely involve reference to other as yet undefined terms. The recursion of definitional reliance on undefined terms is the central condition of logic.

If we now add c to a + b, we can make another stipulation.

$$(a + b) + c = a + (b + c)$$

This is **the axiom of associativity.**

Finally, we can state that if we add any two numbers together, we will always get a third number:

$$a + b = c$$

This is **the axiom of closure**.

But we now run into an ambiguity of the kind we mentioned earlier. Unfortunately, there is no consensus on whether the set of natural numbers includes zero. Clearly, if it does the axiom of closure might generate an ambiguity. If a = 3 and b = 0, a + b does not equal a new number c but 3. The question is: Is this a proper instance of counting? To head this possible ambiguity off, we should specify the set of natural numbers more precisely which we can do by formalising our initial definition as follows:

Let the positive integers from 1.......n be the elements of the set of natural numbers ***J.*** Then

$$\forall i, j \; \varepsilon \; \mathbf{J}, \;\; i + j = k$$

In English, this reads "For all i and j which are elements of ***J**** (that is integers between 1 and n):

$$i + j = k$$

Clearly since zero is not included in ***J***, under this definition the axiom of closure does not apply to additions which involve zero.

Using these axioms and treating multiplication and division as recursive addition and subtraction (which would mean having to define recursion formally as well!), we can extend our deductions and definitions to include them as well and consider other kinds of numbers such as the rational and real numbers. Pretty soon we will find where our axioms do and do not apply to the way we want to use these number systems and, therefore, where new axioms will have to be formulated and new definitions provided. Once we have a consistent set of axioms, we can begin formulating and proving *theorems* (or derived propositions) for the number systems we have described.

To recap. Formalisation involves taking propositions expressed by a small number of undefined terms and translating them into a defined notation. Using the rules of use for the notation, axioms which we take to be true (at least for this argument) are derived. With these stipulated-to-be-true statements, we then derive and prove further propositions or theorems based on the axioms together with whatever auxiliary terminology we may choose to introduce and define via the undefined terms we started with. Apart from our undefined terms, everything in the unfolding argument is formally specified and each step laid out. The ordering is one of linear dependency. The theorems or propositions are built up from definitions and axioms by use of the rules of inference.

We have a logic of formalisation: undefined terms -> definitions –> axioms –> deduced theorems….. The purpose of formalising a set of propositions, we assume, is to obtain an improvement in the rigour of the theory. This is because it is not just the proofs, the QEDs of the formal system which are important. Equally important is the manner by which they are obtained. Each step has to be fully derivable from prior steps using strict rules of inference thereby allowing a continuous trace back from theorems or propositions to axioms to a set of clearly laid out definitions. In addition, we might hope for a similar scale change in the scope of generalisation which theories can encompass — after all, even in our simple case we went from an instance of counting strawberries to some principles for using a designated set of natural numbers.

MODELS, FUNCTIONS AND IDEALISATIONS

Formalisation is a logical process which we apply to theoretical propositions. For sociology, these theories describe states of affairs or putative states of affairs and propose relationships among them. Formalisation in empirical disciplines, then, involves a process of phenomenological reduction by which our experience of such states of affairs is translated into abstract, formal propositions. Deductive inference is then applied to these abstractions. Once a formally valid set of theorems or hypotheses has been constructed, these are then mapped back onto the empirical states of affairs with which we started. The question is how we are to get from empirical statements and descriptions to abstract propositions and then back again? That is, how do we reveal the structure of the propositions in the empirical theory so that the theory can be formalised and then apply the formalised theory to the world of experience? In sociology, as in most disciplines, this is done by building models, the vast majority of which are mathematical, and interpreting their formalisms as functions of data or idealisations about phenomena, and sometimes as both.[13]

Functions

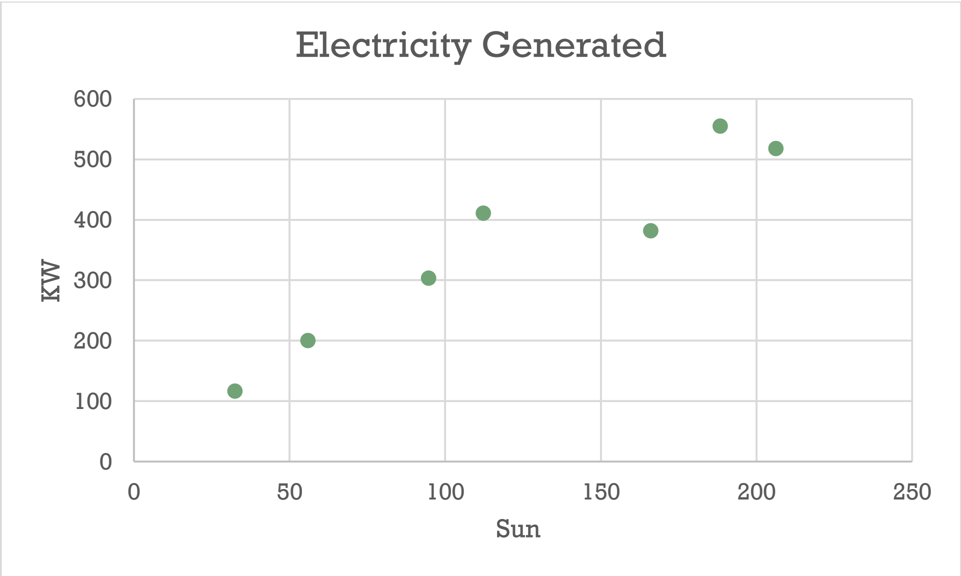
Once again we will use a simple case. Suppose we have installed photovoltaic cells on the roof to generate electricity. Suppose also that for a designated time period we collect data on the amount of electricity generated and count the number of sunshine hours each day. The data gathered is tabulated as follows.

---

13 The best discussion of the application of scientific modelling approaches to social science remains BRODBECK (1958). See also MORGAN AND MORRISON (1999).

Month	Kw	Sun
1	116.4	32.5
2	200	55.9
3	303.4	94.73
4	517.6	206.28
5	411	112.25
6	555	188.36
7	382	166.12

We then draw a chart like this



From the distribution, we find the least squares regression equation for the distribution is

Kw = 2.73 sun hours + 42.2

This equation is a standard linear function: y = mx + c. Just to be pedantic for a moment (it will matter later), the function says that the *variables* x and y stand in the relationship y = f(x) and we have *parameterized* that relationship using the parameters m and c to stand for the slope and intercept. We chose a linear function because the distribution looks as if it approximates to a straight line. To be fully formal the function should be y = mx + c + *e* where *e* stands for the 'error' or distance each point is from the notional straight line. Kw = 2.73 sun + 43.2 + e, then, is a parameterised mathematical model of our distribution interpreted under the formalism of a quantified linear function. Using this abstract model, the algebra of arithmetic and some formal logic, we can theorise about our empirical data and deduce propositions concerning electricity production for sunshine levels we have not had or the level of sunshine for a given level of power generated. We can also infer the likely impact of a combination of factors such as panel size, the length of the cable run back to the inverter, average daily temperature and lots of other things we have disregarded in building the model (they contribute to the error term) but which undoubtedly do have an effect. By designing different set ups to manipulate these conditions, we could look for ways to reduce the modelled error.

Idealisation

When mathematics defines a square as a plane figure with 4 equal sides and all interior angles of 90°, this is an idealisation of the shape of the handkerchief in your pocket, the market square in town, and the beer mat under your glass. All of these may be squares as we know them but the sides may not be of identical length and corners may not be exactly right angles. The mathematical object 'a square' always and necessarily conforms precisely to the definition. Or take the formula for the force of a block moving down an inclined plane.

$$F = f(ma)$$

Here the force of the block is simply a function of its mass and acceleration. This formula describes an idealised block and plane and ignores friction, temperature, air pressure, colour of the block etc. etc., some of which (friction certainly, temperature and air pressure perhaps, but not colour we should think) might well be relevant. As with the electricity example, the model is a linear function and once more using it, the algebra of arithmetic and deductive logic we can make predictions, design tests and so on. We could, of course, use any idealised formalism to model with. Some time ago, there was much interest in modelling social systems with the mathematics of servo-mechanisms. More recently, there has been interest in epidemiological models, network models and models that use non-linear dynamics — so-called 'complexity theory' (Easley and Kleinberg 2010). In each case, what is involved is the adaptation for use in sociology of mathematics developed for quite other purposes.

The value of functional and idealised formalisation, then, is two-fold. First, adherence to the procedures of deductive logic ensures the inferential steps in building the argument are transparent, valid and can be reviewed. Second, the formalism allows us to call on the array of definitions, theorems and related structures of proof which have already been developed within mathematics or logic without having to develop them ourselves. This makes our work a great deal more efficient. Of course, everything turns on the appropriateness of the formalism. If it is inappropriate to the phenomenon we are analysing then no matter how rigorous our deductions, the resulting propositions will be spurious.

Isomorphism

The key issue with functional and idealised models is what is being included and what is being left out during the abstraction process and why. This is the question of isomorphism. How far does the model fit with our description of the characteristic properties of the phenomenon for which it is a model? As you would expect there are clear guidelines for forming isomorphic models.

Let's take yet another homely example. Suppose we took an old pram base and wheels, some planks of wood and string (plus other bits and pieces) and made a go-cart. If we painted it red, put some decals and a number on it and gave it to the grandchildren, no doubt it would give them hours of fun. But no-one, not even the children, would call our go-cart a 'good model' of a Formula 1 Ferrari. Probably the most detailed description of an F1 Ferrari is to be found in the engineering drawings and associated specification documents used in its construction. About the only thing our model has in common with the Ferrari as envisaged in these are the 4 wheels and the red paint. To be an isomorphic model of the racing car as described in the drawings etc, there would have to be a one to one correspondence between the structure of the model and that set out in the drawing; that is, a correspondence of the elements that comprise both and the relations they stand in. In science, for a model to be good this correspondence has to be reasonably complete (though, of course, the goodness and hence completeness of a model is determined by what you want to use it for). Except at the most abstract level, the components making up our cart are nothing like those of the racing car and the relationships they stand in are entirely different was well. The mapping between our construction and the Ferrari does not satisfy the requirements for isomorphism.

On the other hand, our model of power generation is. Physics measures generated power in watts. This, in turn can be defined in terms of two other numerical properties of electricity, volts (V) and amps (I). That is, $W = f(V,I)$. Sunshine hours are also numerical quantities. Both of our variables numerical and so can be interpreted within a formalism ($y = mx + c + e$) which relies upon axioms of the algebra of numbers. The elements or components of the model are, we might say, *formalisation ready*. Equally, the way the model behaves (that is the shape of the distribution) approximately follows the standard linear form. In that sense, the model *complies* with the requirements of the formalisation. The better the isomorphism between model and phenomena, the more ready and more compliant it is and the narrower the gap to be closed when we re-map our formalisation back onto empirical phenomena. The question for any formal model, but particularly those within sociology, is just how narrow that interpretive gap actually is.

Isomorphism is not only concerned with the process of abstraction from empirical phenomena to formalism. It also applies to translations between formalisms. One of the objectives of mathematics is to provide rules by which different formalisms can be translated into each other. For example, alongside the algebra for numbers, there is an equivalent algebra for geometry, matrices, different types of graph, and so on. The existence of isomorphism between the algebras allows us to move between formalisms to analyse our phenomena, always providing, though, we have secured an initial successful formalisation.

If we are successful in translating a sociological theory into a formalism, the one thing we cannot claim is that the theory *qua sociological theory* will have been improved just because it now looks like the theoretical statements of the natural sciences — after all, it is only a translation and so must preserve the sense of that which has been translated. It is true that adopting this much of the Kelvin wedge might ward off some of the aspersions usually thrown at sociology but without the ability to satisfy the other conditions we outlined, looking like a science falls a long way short of being one. The appearance of similarity would be no more than a small gain with regard the marketing or political standing of the discipline.

III. The benefits to sociology of formalization

Real disciplinary benefits will involve changes in the form of theorising or, if the process is generally applied, the character of the discipline itself. For sociology, the following are the ones most often claimed.

Consensus on disciplinary boundaries

A little earlier, we said that formalisation requires precision in the use of terms. This means that propositions have clear definitions of what they apply to, what phenomena they account for and hence what phenomena they do not account for. As the body of accepted theory builds up, so does the associated body of designated phenomena and their properties and relations. Over time, this accumulation of defined terms coheres into a systematic ontology and related metaphysics which characterises the discipline. As the Kelvin aphorism makes clear, the metaphysics of physics incorporates the requirement that phenomena be defined in terms of measurable properties.[14] This metaphysics can be used to demarcate one discipline's phenomena from another's in the sense that we can use the definitions to show how one discipline specifies the properties of a phenomenon differently to another. Clarity here would enable us to avoid confusion when considering a chemist's interests, say, in the experiments of Joseph Priestly and those of a sociologist of science. The interests will be different because the experiments are entirely different for the two disciplines.

Just as important as demarcating what a discipline covers is the consequence of demarcating what falls outside its purview. This is where it is thought sociology might make real gains from formalisation. Because it is hard to draw a boundary around the set of proper topics for sociology, the discipline often appears fragmented and diffuse. Martyn Hammersley (REF) once pointed out that it seems almost anything is or can be the topic of sociological reflection and analysis.

> *Europe, riots, climate control, currency, big society, violence, integration, banks, business, crisis, legacy, the arts, charity, media, justice, religion, politics'……….*

[14] This is the key assumption of the Kelvin wedge. Thus, physics is able to describe and theorise about the auditory signal of a Bach cantata but eschews an interest in the pleasure which listening to it gives. Conformity with the Kelvin wedge would require sociology to exorcise all phenomena which could not be defined in terms of quantified properties.

All featured as the leading items for discussion at one recent meeting of the British Sociological Association. Moreover

> *….the titles and abstracts confirmed this diversity, with papers dealing with topics from psychoanalysis and sociology to Polish male adolescents in English secondary schools, and from public communication of neuroscience to religion, democracy and the Arab spring….it seems now that we have a sociology of just about everything.*

Equally, clarity over the phenomena for sociology, how it constitutes social life, would ultimately have to include how its explanations of those phenomena relate to the explanations of other disciplines which constitute those 'same phenomena' in different ways. What is the proper relationship between sociological and psychological theories of learning, say, or between economic and sociological theories of entrepreneurship, sociological and philosophical theories of the state, and so on? [15]

ADEQUATE COMPLEXITY IN THEORISING

One of the benefits claimed for a resolution to the current theoretical disarray will be the ability to match the complexity of sociological explanations to the complexity of social phenomena. To use a common image, because of sociology's inability to tie its explanations together into a coherent whole even the current variety of sociological topics and sociologies fails to mirror the multi-faceted nature of social life. Hubert M. Blalock, never one to mince words, put it like this.

> *Nearly all sociologists give lip service to the notion that social reality is highly complex. Yet we seem to prefer relatively simple theoretical explanations of this reality, together with data analyses and measurement-error assumptions that are also highly simplistic. Why this is so, and what the implications of sidestepping theoretical and methodological complications in terms of knowledge accumulation of the discipline. I believe that there are a host of exceedingly important questions we need to ask if, indeed, sociology is to advance very far beyond its present confused state. (BLALOCK 1994 P. 121)*

The obvious mismatch between the complexity of the social world and the simplicity of sociological explanations of the parts it studies is precisely what has led to the lowly esteem which sociological knowledge is accorded.

This is not just a demand for complexity in theoretical organisation though — after all, many would claim that the theories of Parsons, Heidegger, Levi Strauss, Levinas, Habermas, Deleuze and so on are quite complex enough already — but for that complexity to be built up through the accumulation of a myriad of less complex (in Kelvin's case, quantified) formulations which

---

15 In these cases the usual stories seem to revolve around tricky matters of epistemic reduction or conceptual supervenience.

demonstrably map onto features of the social world. The hope is that the complexity of the world might be revealed step by step as our understanding of its components is accumulated; each more complex theorisation being formally derivable from those which have already been proven and the new definitions and axioms which research suggests. Achieving adequate complexity in the way which Blalock requires will be no easy task and will require a significant shift in the way topics for investigation are selected. Instead of prioritising topics because they are deemed important by significant vested interests, personally interesting, represent pressing social problems, or because the chattering classes are taken by them, topics will have to be selected on the basis that their investigation is now possible and necessary *given* our current understanding of the social world. To repeat an observation made by Alfred Schutz (1962), in scientific disciplines investigative problems are set by the discipline not by the practitioner. They push themselves forward because (a) they are the obvious next steps to take in the build-up of adequate complexity; and (b) the discipline has the wherewithal to investigate them in ways that will add to the evolution of adequate complexity. In addition, there is Kuhn's key point namely that the discipline licenses the expectation that the problem can be solved. Stanley Leiberson indicates how far sociology is from this, when he points out this will mean the discipline institutionalising a self-denying ordnance whereby research projects which are fascinating but currently undoable are ruled out.

> *The typical project involves working with data sets that are far from ideal. Because of this, perhaps we are apt to avoid the issue of whether* ***some*** *questions are simply not answerable with the tools of empirical social science even if the available data are of exceptionally high quality…….Of course, one cannot simply mutter "undoable" when a difficult obstacle is encountered, turn off the computer and look in the want ads for a new job — or at least a new task. Instead it means considering if there is some inherent logical reason or sociological force that makes certain empirical questions unanswerable. There are four types of undoable questions to consider: those that are* ***inherently*** *impossible; those that are* ***premature****; those that are* ***overly complicated****; and those that theoretical and empirical knowledge have* ***nullified.*** *(Leiberson 1985 pp 7-8 emphasis in original)*

For Leiberson, purging the discipline of such research would go a long way to remedying a defect wherein much of sociology…

> *….incorporates indefensible and illogical procedures accompanied by a certain scientific ritualism. (ibid p13)*

BESPOKE FORMALISMS

Naturally, Leiberson is not proposing to give up on formalisation. He believes that done properly (properly by his lights, that is) it will deliver benefits. The trouble is that there are many formalisms we could use and it is by no means clear which, if any, are best suited for social phenomena. One of the perennial complaints of those like Blalock and Leiberson who campaign for *more* mathematical formalisation in sociology, is that most of the mathematics deployed in it at present is either ill-used or abused. The widespread, almost institutionalised, mangling of what are mainly fairly basic statistical analyses is the leading case in point.[16] But, even if mathematical formalisms were used as they were designed, it is pretty clear that most of the options on offer at present are far too simplistic and clumsy to be of any extended analytic value and so hardly demonstrate the value formalisation might have in facilitating enhancement in the discipline. As some mathematicians have themselves observed, what is needed is a mathematics designed and developed to provide the rigour of abstraction for social phenomena that arithmetic, geometry, trigonometry, topology and the rest do for physical objects and space.

> *Many mathematicians have the impression that mathematical problems in the social sciences are entirely trivial. On the contrary, most problems in the social sciences are too difficult for present-day mathematics. It is because the problems arising in the social sciences rapidly become difficult that only some of the very simplest mathematical problems have been solved so far. (Kemeny 1959, p.578)*

The point that Kemeny is making is one which is often underappreciated by those who promote the adoption of modes of formalisation simply because they are used in the natural sciences. In the natural sciences, formalisation has been worked out in tandem with the working out of the science it serves. The two go hand in hand. When mathematics is used, it gets its purchase on the phenomena studied because the formalisms and the conceptualisations are closely correlated. This intimacy is precisely what is missing in the social sciences.

SYSTEMATIC OPERATIONALISATION

Improvement in the forming and use of theory are not only expected with regard to the process of abstraction. The steps to map conceptual abstraction onto empirical phenomena will be formally secured as well. The natural sciences do this through highly specified systems of measurement which have been developed alongside the particular metaphysics which each science has elaborated. Using such metaphysics, every science has created its own methods, devices and measures. As we have just seen, there is general agreement sociology has no clear sense of what limits there are on the range of its phenomena nor how they are defined and related. Further, although it does have a distinctive common terminology and has introduced a host of supposedly technical terms, as James Davis (Davis 1994) pointed out, these are rarely re-used other than by their originator. As a consequence, sociology ends up

16 The ESRC Review Panel we mentioned at the start was concerned about the degree of ignorance of statistics within the UK sociology profession. We would add the prevalence of inappropriate or misused statistics when they are used.

relying mainly on natural language to gather its data and communicate its ideas. The meaning of sociological terms is enmeshed with the meaning given to concepts used in ordinary life; meanings which, as we have already noted, are notoriously informal, discursive and reflexive. To get their data, physicists, biologists and geomorphologists have the luxury of not having to depend on their phenomena's understanding of the questions being posed by investigations and so can use conceptual formal and quantitative modalities. Unfortunately (or fortunately), sociologists do rely on their subjects to inform them and so must use a modality with which their subjects can engage. Natural science's data gathering instruments can be rigid and precise; sociology's have to be flexible and imprecise. Aaron Cicourel (Cicourel 1964) pointed out that to get the precision of science, we would have to find a way of instrumenting ordinary speech so that the statements our informants make can be translated into formal quantified modalities we could deploy.

Rigorous procedures for formalising theory and equally rigorous methods for mapping formal theory back onto the empirical world will be needed to derive the benefits hoped for from formalisation. *Prima facie b*oth look to present significant challenges. Unless we overcome these challenges, we will be unlikely to fulfil the requirements of the Hypothetico-Deductive and other schemas often said to be the strategies for generalisation used by the natural sciences. Decades of experience have shown how hard it is to cast sociological theory into forms amenable to these schemes without the resulting propositions amounting to little more than tautologies. This implies yet another task; the creation of a new mode of formal generalisation whose requirements sociology could satisfy. Alas, there seems to be no sign of that in the offing either.

In this and the previous section, we have looked first at what is required for a discursive, informal theory to be formalised and some of the benefits which might be gained from doing so. These benefits were held to be (a) improvements in the character of the theoretical propositions being formalised and (b) improvements to the nature of sociology as a discipline when the general standard of theory had thereby been raised. In the next three sections we will look at examples of the formalisation of social theory. Each brings out a different aspect of the challenges we have sketched. All three indicate just how difficult it will be to achieve formalisation of the core of sociology as it currently is carried on and just what cost attempts to impose formalisation would incur.

IV. Formalising a theory of organisations

In this section we simply look at the challenge of translating an existing informal sociological theory into a formal notation. The example we consider is Kamps' (Kamps 2000) formalisation of J.D.

Thompson's (THOMPSON 1967) highly respected theory of organisation. The purpose is to show the nature of the task set in trying to establish the integrity of a set of formal propositions extricated from the informal statements in the theory.[17]

We used the term "extricated" advisedly. Kamps does not just take sentences and phrases from Thompson's text and arrange them in a clear structure. He has to do a fair amount of transformative and creative work to patch the statements together sufficiently to enable formalisation. The conceptual lacunae, missed steps, contradictions, non-theorised components and background assumptions endemic in all informal theory have to be explicated, remedied or rectified. To give an idea of what this work entails, we will look in some detail at how Kamps goes about proving the very first of Thompson's theorems.

> *"Under norms of rationality, organisations seek to seal off their core technologies from environmental influences". (Thompson, 1967, p15)*

To prove this theorem, Kamps works through a number of steps.

Step 1.

Although Thompson bills his theory as a theory of complex organisations, at no point does he actually define what complex organisations are. Taking "organisation" and "sub-organisation" as undefined terms, Kamps defines complex organisations as organisations containing internal organisations which focus on the technical (that is the manufacturing or service delivery) aspects of the organisation's performance. Using the labels **O** for organisation, **SO** for suborganisation and **TC** for technical core and the notation of first order logic, Kemps axiomatises a complex organisations as:

a) Axiom 1

$$\forall x\, [\mathbf{CO}(x) <\!\!-\!\!> \mathbf{O}(x) \wedge \exists y\, [\mathbf{SO}(x,y) \wedge \mathbf{TC}(x,y\,]]$$

In English this reads "X is a complex organisation if and only if x is an organisation and there exists a y such that y is a suborganisation of x and y is the technical core of x".

Thompson does not explicitly say that a complex organisation only has one technical core although he repeatedly talks about "the" technical core. Using the above definition, Kamps goes on to specify the axiom about the uniqueness of the technical core.

b) Axiom 2

$$\forall x,y,z\, [\mathbf{TC}(x,y) \wedge \mathbf{TC}\,(x,z) \text{ -> } y \equiv z]$$

In English: "If y and z are technical cores of x, then y is identical to z.

[17] Kamps uses a computational theorem prover for the derivations. This is immaterial to our interests though it does reflect some considerations with regard to the surveyability and scrutiny of formal systems we will raise later.

Step 2.

The theorem uses the concept "norms of rationality". Again, this term is not precisely defined. Kamps interprets it as "technical rationality", that is, the shaping of means to ends. The use of technical rationality to assess the state of the technical core, he labels **REVA**. Thus **REVA**(o, tc) means that o evaluates tc in terms of technical rationality. This allows a further specification:

$$\forall x \in \mathbf{CO} \rightarrow [\mathbf{REVA}(x,tc)]$$

In English; "If x is a complex organisation then its core technologies are rationally evaluated"; that is, evaluated in terms of means/ends efficiency.

Using this axiom, Kamps then defines the purpose of **REVA.** Having evaluated the technical core, the organisation will seek to reduce any uncertainty affecting that technical core. This is expressed as **RED**(o,u,tc) where u stands for the level of uncertainty in tc.

c) Axiom 3

$$\forall x,y,z\ [\mathbf{SO}(x,y) \wedge \mathbf{REVA}\ (x,y) \wedge \mathbf{UC}\ (y,z) \rightarrow \mathbf{RED}\ (x,y,z)]$$

In English: "If y is a suborganisation of x and x rationally evaluates y and z is the uncertainty of y, then x attempts to reduce uncertainty z for y".

Step 3.

Kamps proceeds in much the same way with regard to the other unexplicated part of the theorem, the sealing off of the core from uncertainty generated by environmental influences. Axioms about the existence of environmental influences and the uncertainty they generate are provided. The notion of "sealing off" is **RED** as applied to the special case of environmental factors. The term **SEFF** (o,i,tc) means o seals off tc from i. This enables a fourth axiom to be defined.

d) Axiom 4

$$\forall x,y,z\ [\mathbf{SEFF}(x,y,z) \leftrightarrow \mathbf{SO}(x,z) \wedge \exists v,w\ [\mathbf{ENVI}(z,y,v) \wedge \mathbf{UC}\ (x,y) \wedge \mathbf{UC}\ (z,w) \wedge \mathbf{C}(y,w) \rightarrow \mathbf{RED}\ (x,wz)]$$

In English: "x seals z off from y if and only if z is a suborganisation of x, and there exists v and w such that z is exposed to influence y from v, and y causes uncertainty w for z, and x attempts to reduce w in z.

With these axioms in place, Kamps uses a computational theorem prover to show the initial theorem can be derived from the definitions and assumptions just given.[18]

[18] Thompson's axiom takes one line of prose. Kamps' proof occupies tens of pages. And even then he presents an abbreviated form. What this might imply for the character of formalised sociological theory is something we will come back to.

Discussion

Kamps' exercise was a re-writing of Thompson's theory in first order logic to show how it could be made more rigorous and transparent. The regimentation which results is certainly more formally rigorous than Thompson's original. It brings out many unstated assumptions, undefined terms and unspecified relationships. By means of his tortuous annotation and amendment, Kamps rectifies some of the issues thrown up. In that sense, we can say Kamps clarifies what is clarifiable in Thompson's theory. However, the order of vagueness remaining does not derive solely from the use of a small number of undefined terms. It points to a first and very important consideration. It is in the very nature of informal generalised abstractions that they can't be fully spelt out and so any translation of them must be indeterminate. Expressed as they are in natural language, their meaning is relative to the context in which they are given. The theory's propositions say that organisations will act in this way or that but they do not and cannot exhaustively set out what is to count as instances of this or that way of acting. In the original theory, these statements work because they are offered under the operation of what Harold Garfinkel (Garfinkel 1967) called "the *et cetera* clause" where, in ordinary usage, the enumeration of any concept's meaning carries the subscript "…and so on". In Thompson's propositions, readers were expected to determine for themselves how far the examples they were concerned with fit under the scheme. The meanings of specific concepts in Thompson's theory cannot be fully laid out in closed propositional form but remain open and are worked out in the context of examples used to reveal what they cover and what their limits are.

The implication is one Herbert Blumer (Blumer 1956) drew many years ago. At the time, much play was being made of the use of what Blumer called "definitive concepts" to make theory and related investigations more rigorous. Such definitive concepts, it was claimed, made it easy to decide if any particular case was or was not an instantiation of the concept. However, as far as Blumer could see, the concepts were actually more "sensitising" than definitive. As such, they provided "a general sense of reference and guidance in approaching empirical instances". Rather than providing a prescription or specification against which any instance could be evaluated in a rule like way, the fit between concept and instance involved weighing a diverse set of considerations. What the concepts really provided were guidelines for finding "family resemblances" among phenomena rather than definitional identification.

The point is not that it is impossible to form definitive concepts to operate in a rule like way but that the variety of forms through which concepts are instantiated in sociology is combined with the relative lack of empirical detail which sociologists hold on those forms, thereby making the construction of definitive concepts extremely difficult. A concept such as "complex organisation" does not define a species of organisation but points to the myriad of ways in which prisons, hospitals, schools, or industrial plants are somewhat alike. As each new case is brought forward, a new endeavour being investigated for example, assessment of fit to the concept is made by interpreting the concept against the case and the case against the concept. This results in a degree of looseness or play in the relationship between the concept and the instance. The degree of play is usually not so material that it prevents sociologists understanding the rough point being made. On the other hand, is too inexact to be definitive.

In Kamps' reformulation, the assumptions, definitions and propositions of Thompson's theory are formally arranged.[19] But, as we have observed, many of the terms are not and cannot be formally defined. While formalisation might have made the theory somewhat more logically transparent and deductively secure, the translation retains an indeterminacy which still has to be resolved on a case by case basis. At the same time there is an underdetermination of the theory's conceptual structure. We have a set of axioms and a theorem describing properties of complex organisations. Such organisations are distinguished by possession of a sub-organisation focussed on the technical core. It is easy to see how this definition fits organisations whose business is the production and sales of goods and services. But how is it to be extended to organisations like prisons, hospitals, major world religions and the like which sociology treats as every bit as complex but where it would be hard to designate a technical core? Equally, the definition states there is one and only one technical core. What is the basis for this? Conglomerates like the British Airports Authority, Boeing, Samsung and Rolls Royce could be said to have many technical cores since they are bundles of correlated businesses. Just how would we be expected to map these organisations into the formal definition? Again, in ordinary usage, 'organisation' sometimes means 'company' and sometimes means 'plant' (among other things). The implications of this cluster-like nature of the meaning of the term are not taken up by either Thompson or Kamps.

Similar questions arise with regard to extending the definition of "norms of rationality" to the array of decisions and actions which any organisation displays. Kamps defines these norms as the reduction of uncertainty through the adjustment of means to ensure the attainment of ends. We can be fairly sure some organisations sometimes do take action in this way. Companies build up inventory, seek to manage their supply chains and tie in their customers. They also try to access R&D innovations, partner with other companies and buy out their competitors. All of these could be called sealing off or protecting their core in order to manage the uncertainty generated by environmental factors. But is this the only rational strategy organisations adopt and is it adopted throughout the organisation? After all, companies sometimes regularly give up a line of business they decide they can't compete in, thereby adjusting ends to means. Yet again, they sometimes persist in a line of business even though they know they are uncompetitive and losing both customers and suppliers simply because they have no other options 'on the table' at that point. Equally, if we accept the priorities, interests and relevances of the organisation are not only those defined and endorsed by senior management, then there is no doubt when we look on the shop floor, in the backroom offices and out among the sales force, we will find many other ways in which norms of rationality are defined.

[19] This re-arranging is so thorough going that we might be forgiven for suggesting that the formalised theory fed into the theorem prover is more Kamps' theory of organisations than Thompson's.

The regimenting of Thompson's theory has demonstrated that some limited formalisation is possible. Nonetheless, indeterminacy of translation and under determination of conceptual scope remains. Given the current practice of sociology, the translation is not and cannot be fully formal. For us, this implies that attempts to reverse engineer formalisation for extant theories is likely to mean the degree of formalisation achieved falls short of the standards required by formal disciplines. This being so, the original intent of the exercise must be cast into doubt.

V. Identifying Causal Chains

From the outset, causal accounts of social phenomena have been central to sociology. Unfortunately, no consensus exists on how these should be formed, nor indeed on whether they are appropriate for sociology's topics. Recently, Judea Pearl has developed a formal semantics which, with appropriate adjustments, many hope will can be applied to offer a procedure for tracing and assessing the relationships set out in sociology's causal explanations.[20]

The paradigm example of causal inference is the one Hume used. When we see a moving billiard ball collide with a stationary one, we *infer* that the momentum of the first causes the displacement of the second. Putting it more precisely, we infer a causal relationship across the following *propositions*.

1. Ball A travelling at speed s collides with ball B.
2. Subsequently ball B is displaced distance d.
3. Therefore: ball A caused the displacement of ball B.

Should we wish to, we could easily imagine instrumenting the table and the balls and so measuring the velocities, displacement and forces involved. Finally, having repeatedly rolled one ball into the other while varying the speed of propulsion, the masses of the balls, the surface of the table etc. and taken the measurements, we could expect to be able to specify a mathematical function across the relevant variables. This, at least in very simple outline, is the way the natural sciences arrive at the rigorous causal explanations they offer. True, as the problems get more complex, the instrumentation gets more complicated, so the mathematics gets more elaborated and the causal chain gets longer. But the shape of the overall story remains the same. By measuring and manipulating the variables in the model, we come to see how the causal process works.

20 The key resource is PEARL (2009A). This is a technical presentation and covers the theoretical background to the semantics. Other less technical introductions are PEARL (2009B AND 2010) and ELWERT (2013)

Standard accounts of the concept of cause used in science (and hence of the efficacy of its causal explanations) insist the following conditions:

1. A causes B iff A temporally precedes B.
2. A causes B iff A is local to B.
3. A causes B iff there exists a mechanism, M, such that A in the presence of M produces B.
4. *Ceteris paribus,* given 1, 2, 3 above, the caused result *always* occurs.

This is the Laplacean universe of classical physics. For some effects described by Quantum Theory, conditions 1, 2, and 4 appear to be violated. The manipulation of condition 3 is what underpins the interventionist method of the controlled experiment. The *ceteris paribus* caveat in 4 will be important at various junctures in our discussion.

The problem for sociology is that we seem unable to determine if our explanations satisfy the above conditions. The main methodological (as opposed to conceptual) reason offered is that the nature of the causal processes in which sociology is interested largely prevent direct manipulation of the social environments of individuals or groups.[21] Ethically, we cannot undertake randomised controlled experiments to vary the ordering, timing or relationships of the most significant variables associated with individual's social experience. Even if we could, we cannot assume we can hold constant all other variables except that manipulated by the intervention. That is, in changing the value of an individual variable, we cannot be sure we do not change the values of other variables defined for the function being examined?[22] The consequence of our inability to exercise control over the set of variables we are studying is that we cannot be certain we have specified *all* the relevant variables and determined their *precise* ordering. This is called 'the identification problem' and represents a challenge to the isomorphism between our models and the social systems we study.

Although controlled or other 'experiments' are not available to us,[23] by means of surveys and the collection of data from official and other sources, we can obtain statistical measures of association from which we can build structural equation models (SEMs). These can then be visualised as causal path diagrams based on directed graphs. The arrows or edges between the nodes in the diagram indicate the strength and direction of theorised association.[24] These techniques were introduced into sociology by

---

21 For the moment, we are going to disregard the question of the fit between Laplacean assumptions and social science concepts of social action, though we will return to it briefly later.

22 This is the key SUTVA (Stable Unit of Assessment Value) assumption which underpins all intervention based methods including the counterfactual interventions which Pearl advocates. As we will see, it poses a severe problem for the validity of interventionist methods in sociology. A second assumption which may give us pause is the assumption of 'population' heterogeneity for any level of causal explanation. Can we be sure that we have pitched the explanation at the level where differences across individuals make no difference to individual causal outcomes?

23 We should note *en passant* that not all sociologists agree. As we will see in the next section, the whole rationale of NET and Elementary Theory is to generate just such experiments. We have discussed the issue of experiments in social science elsewhere. See ANDERSON AND SHARROCK 2014.

24 Since every beginning student of statistics learning to chant the platitude "Association is not causation.", it is clear that what Path Analysis means by "cause" cannot be the same as Hume's meaning since this use cannot satisfy the conditions enumerated earlier.

Blalock and Duncan in the 1960s (Blalock 1985, Duncan 1975).[25] Over the next several decades, structural equation modelling of this kind came to dominate many areas of sociological investigation. Take up was reinforced by the availability of easy to use computational tools. By the 1990s though, concern was being raised about the rigour of the studies carried out. Simply because investigators could compute beta coefficients for large numbers of variables and map them as path diagrams did not, of itself, mean that the underlying theory specifying the variables was causally robust. Indeed, enthusiasm for path analysis seemed to be leading to fewer and fewer credible fine grained 'causal explanations' of more and more social processes.

Counterfactual Analysis

The dissatisfaction with SEMS-based Path Analysis has led many to return to the original impasse to causal investigations in sociology, namely the inappropriateness of interventionist or experimental set ups, and to seek an equivalent methodology which could be deployed within sociology. Pearl's counterfactual semantics for causality seems to offer this possibility. This semantics is based upon the potential outcomes method of statistical reasoning developed for biology by Neyman (Neyman 1935) early in the last century and the logic of counterfactuals developed later by David Lewis (Lewis 1973a, 1973b). Both were extended to the social sciences by Donald Rubin (Rubin 2005). Although Lewis certainly was not interested in practical applications of the logic of counterfactuals within empirical disciplines and Rubin had mainly statistical considerations in mind, Pearl himself does allude to applications in 'policy sciences' like social welfare economics. It comes as no surprise, then, that scholars such as Stephen Morgan and Christopher Winship (Winship & Morgan 2007, 2012) and Felix Elwert (Elwert 2013) hold considerable hope that counterfactuals may be a way of addressing the identification problem and thus of making causal explanations in sociology more robust. Such hope, though, is mixed with caution about the gap between what sociology currently can do and what counterfactual causal semantics requires it to do. Everyone is well aware there is still a long, long way to go. Morgan and Winship talk of the many challenges faced as the consequence of being "on the frontier" of research. And as we will see, they are right. Counterfactual causal analysis is pioneering stuff. It will be a long time before its use will be mapped, regulated and settled such that we can realistically evaluate what, if any, *improvement* to sociological theory might be gained from it.[26]

Strengthening the robustness of causal explanations in the social sciences was not the immediate objective for the mathematical machinery Pearl and his associates constructed. Even today, it remains off the main line of their concerns. These are based in the general AI programme for formalising human reasoning and thus facilitating its automation. One of the central challenges to *that* programme

---

[25] See Lleras (2005) provides an introduction to Path Analysis and Morgan (2013) gives a summary of the history of causal modelling in sociology and the similarities and differences between Blalock and Duncan's positions over the 'reality' of these causes. Russo (2009) and Russo et al (2010) give some detail on the philosophical issues.

[26] One of the many important aspects we will not discuss here is the body of assumptions and adjustments which have to be made in regard to the sociological data that has been gathered for it to be suitable for counterfactual analysis. (See Morgan & Winship 2007 p 21.)

(if not *the* central challenge) is that much human reasoning is non-deterministic. We seem to be willing to grant the truth of a proposition even though that proposition cannot be derived logically from other propositions. This is particularly so for causal relationships. To use the example we started with, we infer the force of the first ball that displaced the second ball even though there is no entailment between the propositions describing their states. As Hume put it when he introduced the story, our inference of a causal relationship between the states is not a logical response, but a psychological matter of predisposition, habit or custom.

Pearl accepts Hume's point and asks if there could be another, weaker basis on which the logic of causal connectivity might be formally framed. He proposes that, appropriately organised, the calculus of probabilities combined with the semantics of directed acyclic graphs (DAGs) will suffice. The appropriate organisation is contained in his sets of axioms and derived theorems together with the proofs offered for them.[27] The reasons for choosing probabilities are twofold. First, the 'language' of probabilities is what the majority of relevant domains in statistics utilise. Second, causal thinking in natural language is probabilistic too. When we say that setting the oven temperature too high caused the cake to burn, we know that on another day the cake might well burn even though the temperature was correctly set (we left them in the oven too long?). So, when, the following weekend, we are faced with yet another burnt cake we have to determine the probability that the oven was too high, the time too long or something else went wrong (we are just inept at making cakes?).

The probability structures Pearl uses are Bayesian. That is, they represent beliefs about the distribution of a set of variables, beliefs which might change as we gain knowledge about them. This is an important break with previous approaches, at least to causal explanation in the social sciences. Pearl postulates that causal explanations are subjective beliefs couched as non-formal causal theories. Thus, by extension, causal explanations in sociology describe the beliefs the investigator holds as to the causal processes at work.[28] The aim is to find a way of testing the transparency and logic of these beliefs.[29] Thus if A stands for the proposition "It will rain this afternoon", $P(A \mid K)$ can stand for our subjective belief that given what we take ourselves to know, it will rain this afternoon. Conventionally, this notation is abbreviated to $P(A)$. Clearly, as we acquire (or lose) knowledge, the value of $P(A \mid K)$ will change. The revision of subjective belief and the change in P values associated with it are the hook on which the proposed method of counterfactual intervention is hung. Bayes Rule says that the overall strength of belief in $P(A)$ is the sum of the odds ratio (the frequentist predictive belief $P(A)$) and the likelihood ratio (diagnostic or retrospective belief) that A occurred because of B ($P(A \mid B)$). Pearl's method changes the likelihood ratio by intervening ('surgically', as he puts it) on B in the schema and

[27] The work involved in providing this organisation should not be underestimated. The scale of Pearl's achievement has been widely celebrated in the mathematical literature. Some, though, remain doubtful about his programme (Cartwright 2007a)

[28] This is important. It opens up the possibility of a potential mis-mapping between the investigator's belief set and the actual empirical causal processes. Later, we will see this potential is critical for sociology.

[29] This sets one of the key challenges for Pearl. His formalisation has to show the mathematical equivalence of Bayesian and frequentist associational distributions of a system of variables.

conditioning or changing it thereby impacting the other theoretically relevant variables. This intervention changes the value of B and thus the likelihood ratio for A thereby creating the counterfactual. If the likelihood that B causes A changes in line with the intervention, the strength of our overall belief in the causal story will be confirmed. If not, the case for a causal connection is weakened. A critical axiom here is that because the original SEM is in equilibrium, the members of the family of models produced by conditioning are formally equivalent. Adjustment of one variable will generate adjustments across the other variables to maintain the same equilibrium value. The scale of the change of A given the scale of the change on B is the test of its causal force. The conditioning strategies are the various "counterfactual interventions" Pearl develops (e.g. d-separation, instrumentation, back door and front door controls etc.). Changing A to some different value is what Pearl calls the Do(.) operator. Conditioning and the Do(.) operator are the heart of his semantics.

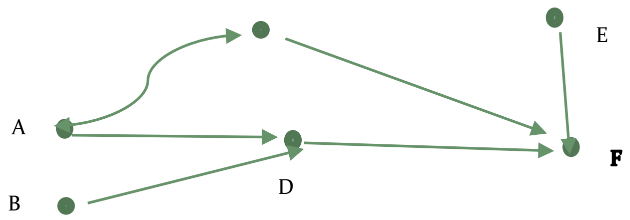
Pearl believes that when they were first developed, SEMs and the path analyses based on them were causal in conception. Over time, because of a professional commitment to a probabilistic interpretation of co-variance and regression in statistics this usage has been lost. He wants to re-instate it.[30]

> *Thus, to the often asked question, "Under what conditions can we give causal interpretation to structural coefficients?" Wright and Haavelmo would have answered, "Always!" According to the founding fathers of SEM, the conditions that make the equation $y = \beta x + \varepsilon$ structural is precisely the claim that the causal connection between X and Y is and nothing about the statistical relationship between x and can ever change this interpretation of β. Amazingly, this basic understanding of SEM has all but disappeared from the literature, leaving modern econometricians and social scientists in a quandary over β. (Pearl, 2009, p135)*

The semantics which Pearl uses is an extended form of the directed graphs (DAGs) used within Path Analysis. Instead of representing the association between variables, the arrows or edges between the nodes indicate the direction of causal influence. The wavy double headed arrow between A and C indicates the existence of common unspecified (or confounding) causes for these two variables. Because there are no recursive paths between variables, the graph is *acyclic*.

C

30 Pearl's interpretation of the history of SEMs and Path Analysis is by no means uncontested. See RUSSO (2010) and MORGAN & WINSHIP (2007).



Each variable represents a Bayesian probability distribution of known and unknown (i.e. implicit) causes. Thus D is the outcome of the causal effects of A and B together with an unknown set of other causes. The relationships form a Bayesian Network. These are the formal mathematical objects Pearl uses. The justification for treating them causally is, first, that the formal equivalence between the original SEMs and their Bayesian interpretation can be demonstrated and second that this is a natural way of talking about them. In other words, Pearl views these networks as quasi-deterministic functional models and so a proper formalism for causal analysis.[31]

As we have said, Pearl's strategy is to demonstrate the equivalence of a set of structural models based upon the SEM which has been constructed from empirical data. This set is generated by 'conditioning' on variables and analysing the effects of d-separation, back doors, path blocking, collider variables, each of which is formally defined and axiomatised. As each model is constructed, a family of counterfactuals to the original model of the observed variables is built up by 'surgical intervention' to Do(x) on an identified variable. Because the SEM of the original observed variables is in equilibrium, each new member of the family represents a new equilibrium among the variables.

> *(T)he essential characteristic of structural equations that sets them apart from ordinary mathematical equations is that the former stand not for one but for many sets of equations, each corresponding to a subset of equations taken from the original model. Every such subset represents some hypothetical physical reality that would prevail under a given intervention. If we take the stand that the value of structural equations lies not in summarizing distribution functions but in encoding causal information for predicting the effects of policies ...., it is natural to view such predictions as the proper generalization of structural coefficients. For example, the proper generalization of the coefficient in the linear model M would be the answer to the control query, "What would be the change in the expected value of Y if we were to intervene and change the value of Z from z to z + 1?", which is different, of course, from the observational query, "What would be the difference in the expected value of Y if we were to find Z at level z + 1 instead of level z?" (Op. Cit. p 158)*

[31] Just to clarify. The causes specified in the model are deterministic. But not all the causes influencing the outcome have been nominated in the model. Unspecified causes have unknown levels of influence on the outcome. The value of the outcome is known, but what proportion of this value is contributed by the known and unknown causes is a matter of belief. Hence the distributions are deterministic and probabilistic.

Each new model allows us to test the causal relationships among the variables. When the value of a variable is adjusted or the variable's path structure is re-shaped, the value of the dependent variable changes, the strength of the causal connection between them is displayed. If the path is blocked but the dependent variable's value remains unchanged then the causal connection is invalidated.

Because they represent different instances of equilibrium among the variables, Pearl holds the approach refutes a standard statistical objection to counterfactual analysis, namely that counterfactuals are "metaphysical" in some Logical-Positivist sense. Pearl's argument goes like this. If we want to test the effect of a treatment where the outcomes are binary for the patients (died/ not died), one question we might want to answer is whether the patients would have died anyway without the treatment (p(D) | no T). Since all the patients were treated, we have no untreated test cases to assess. The standard view is that our question is impossible to answer. For Pearl, this is not so.

> *In our example, the response to treatment of each (surviving) patient is assumed to be persistent. If the outcome Y were a reversible condition, rather than death, then the counterfactual claim would translate directly into predictions about response to future treatments. But even in the case of death, the counterfactual quantity Q implies not merely a speculation about the hypothetical behavior of subjects who died but also a testable claim about surviving untreated subjects under subsequent treatment. ……(It can be shown) that, barring sampling variations, the percentage Q of deceased subjects from the treatment group who would have recovered had they not taken the treatment precisely equals the percentage Q′ of surviving subjects in the nontreatment group who will die if given treatment. Whereas Q is hypothetical, Q′ is unquestionably testable. (Op. Cit. p. 34)*

In Pearl's view, the outcome of this exercise is to shift the meaning of the term 'counterfactual' from the hypothetical to the empirical. For him, counterfactuals are actually short hand predictions. The function *F = f(ma)* is a prediction that at all future times if acceleration and mass are constant then force will be constant and the counterfactual *F' = f(ma')* says the force would be *F'* had the acceleration been a'. These two functions are equivalently empirical because the latter varies the *ceteris paribus* assumption under which *f = ma* operates. In the counterfactual case(s) the value of a has been changed

> *The less obvious answer rests with the ceteris paribus (all else held equal) qualification that accompanies the predictive claim, which is not entirely free of ambiguities. What should be held constant when we change the current in a resistor – the temperature? the laboratory equipment? the time of day? Certainly not the reading on the voltmeter! (Op. Cit. p. 218)*

Pearl's counterfactual strategy requires theoretically well-defined and precisely measured conditions allowing robust specification of the equations of the system (the SEMs). For any predictions to be valid, both the laws at work and the boundary or exogenous conditions must not change. Given these conditions, the DAG semantics allows the identification of the strength of causal connections via the methods of localized surgery.

Before we turn to look at fit between the assumptions embedded in Pearl's semantics and the practice of sociology, there one final observation to make. By definition, no graph is a fully specified causal model. For all variables there are causes which remain implicit. More importantly, common causes across a number of variables may be left implicit. Pearl recognises this means that no matter how sophisticated the mathematical technology used to unpack the relationships, the modelled variables can only be 'sensitising' or 'identifying' ones.

CAUSES, COUNTERFACTUALS AND SOCIOLOGY

Pearl's work has rightly been regarded as a major contribution to Computer Science. Take up within sociology has been somewhat restricted. The two obvious areas of likely interest are social epidemiology and network analysis, both of which use graph analytic techniques. [32] And, indeed, these domains are where counterfactual semantics using DAGs has been most readily assimilated. The number of studies using counterfactuals is growing exponentially. However, even at this early stage, there is disquiet about the possible consequences of an overenthusiastic rush to adopt Pearl's approach without first being quite sure the conditions it requires for its proper use are in place. As we noted earlier, the most important of these is the full definition of an explicit and valid causal model. What Pearl's approach does is test the inferential structure of the graph. A formally valid graph may still lack tight mapping onto the actual causal structure of the phenomena it represents. Many commentators (for example ELWERT (2013) and KNIGHT AND WINSHIP (2013)) fear that, just as with SEMS, the literature is becoming overburdened with highly technical explorations of poorly formed theories leading to errors of causal identification. We noted earlier that the identification problem, that is the selection of an appropriately structured model that is isomorphic with the structure of our experience of the social world, is a non-trivial matter for sociology on both ethical and methodological grounds. For those promoting the use of DAGs and formalisation in general, this is seen as a (difficult but) solvable practical problem. For others, it is more a Leiberson "un-doable" one. Correct identification is, of course, only half the story. For a model to be of practical use, we have also got to provide valid estimation of the causal influences. For many classes of DAG, this may not be feasible.[33] A related issue revolves around the identification of common causes. If we include common causes (as we are likely to want to do in non-experimentally controlled situations), DAGs rapidly escalate to unmanageable proportions with the result that

[32] HOLLAND (2001) gives an overview. For epidemiology, see GREENLAND ET AL (1999) and BERKMAN ET AL (2014). For network analysis see SHALIZI AND THOMAS (2011) and SHARKEY AND ELWERT (2011)

[33] This difference is marked mathematically. DAGs represent non-parametric functions and thus describe non-parametric SEMS, A = f(B, C, D….). No statements are made about the form of the functions (linear, curvilinear, non-linear etc.) or their parameters. We would need such statements if we want precise estimates of the array of causal forces the function stands for.

investigators find it difficult to set out and maintain control over all the relationships at once. As we will see below, this can mean some interventions inadvertently lead to modelling errors.

Just as the operation of formal logic does not question the truth status of the stated axioms but assumes them in order to deduce further valid propositions, DAGs do not question the mapping of the model specified to the reality that has been modelled. Any model is only as good as the specification provided by the researcher. To bring out what this might imply, we will illustrate the kinds of errors that can occur if a model is poorly specified.

Overcontrol

If the causal structure is A -> C -> B, A is an indirect cause of B. If we fail to mark A's indirect relationship on B but condition on C to see what effect it has on B, we block the path from A to B thereby eradicating an indirect causal path. As we mentioned a moment ago, overcontrol is almost bound to be a risk in the large DAGs required for sociological research.

Confounding

In a structure A <- C -> B, A and B are associated only because they have a common cause C. Non-specification of this commonality and consequent failure to condition on C, leaves in place what will appear to be clear, but in fact spurious, association, between A and B.

Endogenous Selection Bias

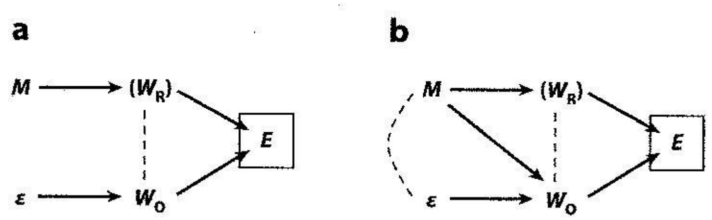
Suppose we are interested in the causes of popularity among undergraduates at our universities.[34] We identify two possible variables, being good at playing the piano (A) and being good at sociology (B). Both are associated with popularity (C). We set up the relationship as A -> C <- B. Here C is a 'collider' of A and B (assuming no causal relationships between A and B and no common cause). We now condition on the collider by filtering the data and only looking at those students who are popular. Under this condition, knowing a student cannot play the piano implies the student is good at sociology, and knowing the student is less good at sociology implies they are good at playing piano. In both cases, we have set up a spurious association between playing the piano and being good at sociology.

Selection bias might seem an obvious and easy thing to avoid, but it is not. Here is an example of what is known as Heckman bias.[35] Studies of the impact of motherhood on women's employment usually gather the following key variables: employment (E), motherhood (M), and offered wages (Wo) for those women who are employed. Typically, data is not gathered on women's reserve wage levels (Wr), the level of wage required to induce them back into employment. Standard theories of employment assume E is a function of both Wr and Wo because a person will only accept a job that

[34] We have adapted this example from ELWERT AND WINSHIP (2014).

[35] See ELWERT & WINSHIP (2014).

offers wages above their Wr threshold. In the case of women, then, the model is M -> Wr -> E <–Wo with E being a collider between motherhood and offered wages (see a below).



Since data for Wo is gathered only for women in employment, investigators generally filter the sample to include only working women. This unblocks the non-causal path from motherhood to offered wages because now there *is* an association between motherhood and offered wages. All the mothers have a value for offered wages (represented as the dashed link between M and ε in b above). The filtering induces endogenous selection bias. The situation is only made worse if we adjust the model to include a direct effect of motherhood on offered wages (through employment practices, or discrimination in hiring, say). We cannot now tell how much of the impact of M on E flows through which of the paths since we are only conditioning on E.

Since they demand the specification of causal variables and their paths, DAGs do provide a technique for surfacing the kind of errors we have just been describing. What they don't do is provide a mechanism for eradicating them from the models we identify.

Discussion

Pearl's semantics is a highly formal procedure for identifying and tracing through causal relationships. As we have suggested, it imposes a number of demands upon the character and structure of the explanation being examined. The question is: Can sociology satisfy the assumptions which the semantics makes about the causal models which it processes?

Assumptions in the Semantics

It is important to remind ourselves of Pearl's objective. He wants to build a 'language' for causal reasoning in AI. This language is to be formal: that is, (a) rule governed; (b) abstract; and in this case (c) mathematical. The language is specified as a somewhat reduced version of natural language. Such reduction is justified as a 'simplification' of the complexity of natural language made necessary in the pursuit of rigor. Whether this search for simplicity in semantics might lead to complication in research, data collection and explanation is something that we will return to.

When fully developed, the semantics will provide a means by which computational systems can evaluate the truth of propositions about causal states. Evaluating the truth of propositions is what Pearl means by reasoning. The function of the semantics is to provide the capacity to…

> *(i) store counterfactual knowledge in the form of counterfactual premises and (ii) derive answers to counterfactual queries using some logical rules of inference capable of taking us from premises to conclusions (OP. CIT. p. 34).*

This objective raises two immediate issues regarding the application of the semantics to sociology. First, the inference engine works on propositions or proposition-like statements about the causal linkages between states. What it manipulates are not states in the world but propositions. While this is obvious, it is not trivial. It means, for example, that when considering any proposed counterfactual, we need to be aware of the disjunctures between model identification, inferential possibility and empirical possibility. Not everything we can infer about a model is possible in the world which the model represents. How infeasible counterfactuals are to be discriminated *within* the semantics is a major challenge. Second, the engine is concerned with beliefs about causal states and only causal states. It is not designed to encompass beliefs about beliefs, desires, values, norms and the like, all of which are important in the construction of causal explanations in sociology. Naturally, propositions about such phenomena could be translated into propositions about causal states (using psychological or other theories, say) but this is likely to do violence to the ways that sociologists want to think about how these interpretive phenomena operate in social life.

In addition, the mathematical character of the formalisation is not incidental and neither is it without consequences. The axiomatised rule structure ensures that proper inferences are truth preserving. However, it does so by excluding reliance on any formulations (knowledge/understanding/inferences) which are not themselves formalised. The reasoner using Pearl's semantics can only reason within the limits of the semantics. His use of the Bayesian probability calculus is an attempt to deal with the well-known and widely discussed 'common sense problem in AI' (Mueller 2006).[36] This refers to the need to resolve 'paradoxes' which are routinely handled in common sense by reliance on non-formal knowledge (contextual, tacit and other); knowledge which everyone agrees resists formalisation. By relying on such knowledge, common sense tells us the following inference is defective. (The example is Pearl's)

i) My neighbor's roof gets wet whenever mine does
ii) If I hose my roof, it will get wet; therefore
iii) My neighbor's roof will get wet whenever I hose my roof.

Pearl's semantics must be able to discover and set aside as irrelevant exceptions, implausible though possible inferences inferred from a set of antecedent conditions. Since, in any real case, there is a potential infinity of such exception generating antecedents, this appears to be a daunting task. In ordinary life, of course, we have no such difficulty since our (informal) common sense tells us that there

[36] A problem which has led many to doubt the feasibility of the programme to which Pearl's semantics is a contribution.

is more than one way to wet a roof without us needing to list in advance all the ways that a roof might be made wet.[37]

In sum, Pearl's semantics is designed for a particular and peculiar usage, formal reasoning by formal systems. This is likely to mean that it and the analytic methods derived from it have a highly restricted domain of application within an empirical discipline such as sociology. Much of the knowledge the sociologist calls into play and much of the data used to undertake sociological analysis falls outside its purview.

Although the semantics is formally constituted, Pearl is well aware the causal models it is deployed upon are not. They are constructed as a matter of (he says "expert") judgment. They are *post hoc* informal rational reconstructions of co-variation relationships between events.

> *.....behind every causal claim there must lie some causal assumption that is not discernable from the joint distribution and, hence, not testable in observational studies. Such assumptions are usually provided by humans, resting on expert judgment. Thus, the way humans organize and communicate experiential knowledge becomes an integral part of the study, for it determines the veracity of the judgments experts are requested to articulate. (Op. Cit. p. 40).*

The semantics just assumes the validity of the model in order to manipulate it. This may be well and good in the case of statements about physical phenomena. Classical physics has a vast store of conventionally accepted propositions about the physical world; a store which has been accumulated through the coordinated and concerted effort of generations of physicists. In sociology, where research is neither coordinated nor concerted, the same does not hold. Almost any set of propositions you might offer about causal states in the social world is open to dispute. Indeed, a not unreasonable description of much practice in sociology is the challenging of other researchers' causal claims. A more forgiving version might be that sociology is still trying to discover what its 'facts' are.

Assumptions about the Semantics

In Pearl's semantics, a model is a mathematical object which describes the (causal) relationships between states. States are defined by variables linked by non-parametric functions. The links between variables represent undefined 'causal mechanisms'. Variables take one of two forms: exogenous variables are determined by factors outside the model being considered; endogenous variables are determined by factors within the model. The semantics of DAGs is only concerned with endogenous variables.

---

37 In sociology, the selection of a few terms for technical deployment as the concepts of a theoretical scheme doesn't reduce reliance on the rest of the natural language but only assigns it to an informal role in theoretical and methodological reasoning. Common sense reinserts itself into the analysis through this use of natural language.

The expression of states in terms of functions brings its own issues concerning feasible and infeasible counterfactuals. Although all variables are formally on the same footing (in that any of them could have been different in some way), in any actual theoretical model not all variables can be functionally different in just any way. Whilst we can readily enough imagine time travel, for example, or teleportation, these are not empirically realizable causal processes which we can allow within any model of a real case under investigation.[38] The 'feasibility' limitations on counterfactuals are set outside the model itself. In like manner, the *directed* nature of the DAG is an attempt to address limitations on the mathematical forms which counterfactuals can take. For example, the expression $f = ma$ tells us that the force of an object is the product of its mass and acceleration. If, for some system, we know the force and the acceleration, using the equation we can calculate the mass. But common sense physics tells us force and acceleration do not cause mass. The causal logic is *from* mass and acceleration *to* force. This might seem an obvious point, but it means that in any domain of application we have to apply disciplinary common sense to ensure the credibility of the causal connections expressed by the formalism. As we hinted just now, disciplinary common sense is often just what is under scrutiny in sociological investigations and agreement on its constituents is far from universal. As a consequence, it seems highly likely that even for the most well ploughed fields of investigation, determination of the limits of infeasibility may be so far beyond us that we would struggle to know how to answer Pearl's admittedly "tough" but necessary questions:

> *What is the empirical content of counterfactual queries? What knowledge is required to answer those queries? How can this knowledge be represented mathematically? (Op. Cit. p.34).*

Moreover, in the models.....

> *...causal relationships are expressed in the form of deterministic, functional equations, and probabilities are introduced through the assumption that certain variables in the equations are unobserved. This reflects Laplace's (1814) conception of natural phenomena, according to which nature's laws are deterministic and randomness surfaces owing merely to our ignorance of the underlying boundary conditions (Op. Cit. p26)*

Pearl accepts this is the metaphysics of classical physics. The problem is that most sociology explicitly departs from this metaphysics. First, given the central concern is *social action* defined in terms of its intrinsic interpretive character, universal *ex cathedra* determinism is ruled out. Social actors as social agents choose their courses of action (even if "in circumstances thrust upon them"). As we have commented a number of times, reconciling social action with the outlook summarized in the epithet "laws of nature" has been and remains a deeply unresolved issue in sociology. Second, explanatory

[38] David Lewis regularly observed that one common response to the examples of counterfactuals he examined was an incredulous stare.

reductionism remains an equally disputed and deeply unresolved issue. The relationships between individuals and society, the macro and the micro and all the other complementary contrasts are still subject to (a) a wide variety of theoretical specifications and (b) an equally wide variety of methods for their investigation. If, as Pearl actually says, Laplacean determinism is a stipulation of his modelling techniques, not only is application to quantum mechanics ruled out (which he is happy to accept), so too is all conventional sociology.[39]

If all this were not enough, the requirement that models be "well defined" is also likely to be just as fatal. We have alluded several times to sociology's inchoate nature. Indeed, the work of defining requisite variables and then robustly evaluating their functions has hardly begun. Sociology has nothing like the panoply of nomological generalisations found in the natural sciences from which functions can be derived. Nor is it likely to have any time soon, if only because a significant part of the discipline thinks such generalisations are of no value and so not worth pursuing. The net result is that, as things stand, we would have a hard time formulating a reasonably complete functional model of some actual state of social affairs to act as the 'control' against which some other, counterfactual 'treatment' based on surgical intervention might be tested. As Pearl explicitly acknowledges, we do not have well defined theories for problems such poverty, intolerance, discrimination and all the other things sociologists are interested in. Until we do, the sophisticated fine grained sifting machinery the semantics provides will not be of much use on sociology's theoretical rubble. Certainly it will not help to adjudicate between alternative sets of theoretical assumptions about causes and their ordering which, at least on the data we are able to collect, appear to have equal theoretical validity. All it will do is allow us to see how far the internal logic of each model is valid.

The febrile character of sociological knowledge has a further important implication. It makes the simplification required to build a model impossible. Pearl thinks such simplification takes place though 'local surgeries' undertaken by domain experts.

> *The world consists of a huge number of autonomous and invariant linkages or mechanisms, each corresponding to a physical process that constrains the behavior of a relatively small group of variables. If we understand how the linkages interact with each other (usually, they simply share variables), then we should also be able to understand what the effect of any given action would be: simply respecify those few mechanisms that are perturbed by the action; then let the mechanisms in the modified assembly interact with one another and see what state will evolve at equilibrium. If the specification is complete .... then a single state will evolve. If the specification is probabilistic.... then a new*

[39] The references to social science in Pearl's work are largely to work in econometrics where, it has to be said, classical mechanics remains the paradigm for model building, and to policy science (or public administration) which appears to have adopted the approach of randomised controlled trials as the gold standard. The robustness of both as ways of rendering the economic and administrative worlds has been subject to much dispute. See MIROWSKI (2002) and CARTWRIGHT (2007)

> *probability distribution will emerge; if the specification is partial .... then a new , partial theory will be created. In all three cases we should be able to answer queries about postaction states of affair, albeit with decreasing levels of precision. (Op. Cit. p. 223-4).*

As he says, what facilities the carving off of the sets of linkages is an appreciation of the locality of action.

> *Standing alone, locality is a vague concept because what is local in one space may not be local in another. A speck of dust, for example, appears extremely diffused in the frequency (or Fourier) representation; conversely, a pure musical tone requires a long stretch of time to be appreciated. Structural semantics emphasizes that actions are local in the space of mechanisms and not in the space of variables or sentences or time slots. For example, tipping the leftmost object in an array of domino tiles does not appear to be "local" in physical space, yet it is quite local in the mechanism domain: only one mechanism is perturbed, the gravitational restoring force that normally keeps that tile in a stable erect position; all other mechanisms remain unaltered, as specified, obedient to the usual equations of physics. Locality makes it easy to specify this action without enumerating all its ramifications. The listener, assuming she shares our understanding of domino physics, can figure out for herself the ramifications of this action, or any action of the type: "tip the ith domino tile to the right." By representing the domain in the form of an assembly of stable mechanisms, we have in fact created an oracle capable of answering queries about the effects of a huge set of actions and action combinations – without our having to explicate those effects. (Op. Cit. p.224)*

The problem is that the sociological equivalent of a 'shared domino physics' is exactly what we don't have and so localization criteria are among the many sources of theoretical and methodological dispute. The lack of localization criteria also makes it impossible to gain agreement across lines of investigation or even within a line of investigation on what are endogenous and what exogenous variables. Indeed, if disputation over causal claims is one way of characterising sociology, then much of that disputation consists in arguments over locality and what is, thereby, ruled in or out of any model.

Where does this leave us? Sociology's unwillingness to endorse a standard strategy of reduction together with the associated lack of agreement on locality criteria mean that it can only solve the identification problem and provide the well-defined causal explanations the semantics requires at the cost of such explanations remaining as shallow as they are today. We all agree that educational attainment is (somehow) a consequence of socio-economic status of household (among other things) but we lack a sufficiently detailed understanding of the causal linkages between the variables to say anything more than that. Moreover, to provide *deep* causal accounts using the counterfactual approach to causal modelling, we would have to sacrifice the interpretive character of social action. The determinist model assumed by the semantics does not allow a step by step relaxation of this stipulation; it requires its excision.

VI. Qualitative Formal Sociology

In our previous examples, formalisation took the form of a reconstruction of an informal theory's deductive structure and the use of a mathematical formalism to assess the integrity of putative causal explanations. The example we consider now is somewhat different. It sets out a formal theoretical argument and uses mathematical formalisms to derive hypotheses or theorems which are used to reproduce analogies of social phenomena. The theory is axiomatic in that it begins with a very small number of defined concepts and sets up postulates on which the theory is built.[40] Given these postulates, propositions concerning the nature of social structures and social relationships are constructed. Derived theorems concerning the forms social structures take and the relationships associated with them are then used systematically to generate analogies of types of social arrangement. The arguments, and hence the formalisms used, are qualitative. They are construed around qualitative rather than quantitative propositions[41]; that is propositions of the kind "x is more or less the same as y", or "x is bigger (or smaller) than y", or "x is (or is not) y (or is almost a y)". Such propositions express either broad comparisons or statements of categorial identity.

The theory is Network Exchange Theory (NET), a component of a broader framework, Elementary Theory (ET), which David Willer and his colleagues have been developing over the last 40 years. (See Willer 1996, 1999, 2003, Willer & Emanuelson 2006, 2008, Willer et al 2012, Willer et al 2013, Walker et al 2000). NET is a species of Exchange Theory, an approach which starts by assuming social action is the exchange of socially valued resources. While ET aims to be a general theory of social structure and social relations, NET has largely focused on exchange structures which predispose the emergence of power relationships.

The name 'Elementary Theory' is not insignificant. The guiding image is that of molecular physics where the elements which compose physical objects are defined as combinations of atoms held together by covalent or other bonds. In turn, atoms are combination of particles bound together by the 4

---

[40] Another way of describing this kind of theorising might be to call it *a prioristic*. From a given set of truths other propositions are derived and tested.

[41] At a pinch you could call them quasi-quantitative as long as it is remembered they fail the requirements of any formal measurement system. They do not represent quantification by the backdoor or 'quantification-lite sociology'.

fundamental forces. On the analogy with physics, society is modelled as structures of social relationships. Structured social relationships are the foundational units. They can be neither reduced to properties of individuals nor do they emerge from them. Rather, as part of the set-up conditions of the model, individuals (actors) are, 'placed' in positions. The model then seeks to show how, given a specified starting arrangement, the structure of relationships will develop.[42] When complete, Elementary Theory will be a general theory of social structural dynamics. It will demonstrate how society 'evolves' by transitioning from one social structure to another. A dynamic theory of social structures will:

a) Provide an account of how a given configuration of relations produces a set of environmental social (i.e. power) conditions. This is what NET aims to do;
b) Provide an account of how the produced conditions lead to new social structural configurations. This is the theory of dynamics.

The formality of NET is provided by a combination of its use of Graph Theory to provide an idealised axiomatization of social structure and derived functional models by which the theory's 'conjectures' are explicated.[43] The use of Graph Theory is unsurprising since graphs of various kinds have been used widely in Exchange Theory and elsewhere to describe networks of relationships. [44]

THE CONCEPTUALISATION OF SOCIAL STRUCTURE AND POWER

Like all Exchange Theories, NET begins with the following generalisation.

> "Social relationships consist in the exchange of socially valued resources or sanctions."

On the basis of this generalisation, the following postulates are stipulated.

1. Conceive the dyad to be simplest form of social relationship and the exchange of positive or negative sanctions as the first principle.[45] The elementary dyad is represented in the following graph.

[42] In a sense, the approach is a kind of *Lego* sociology. From a few simple forms structures are built, the architectures of which are determined by the properties of the basic 'building blocks'. As structural demands become more complex or differentiated, new or morphed building blocks have to be introduced. These retain the same essential 'exchange' properties. The theoretical *desiderata* for the theory become theoretical integrity through the standardisation of essential properties and generativeness through the modes of differentiation.

[43] In early work, explication was by demonstration in experimental set-ups of the usual kind in social psychology. More recently, resort has been made to computationally based simulation. We will leave to one side how far these strategies conform to the principles of Lakatos and Popper's 'falsificationism' to which Willer has consistently claimed adherence.

[44] Many of the early applications of Graph Theory were with social science topics. See Harary (1969) and Barnes and Harary (1983)

[45] Simplicity here should not be taken to mean ontogenetically primeordial but structurally minimal. Exchange theory is not a theory of social evolution.

Although worked out with less sophistication and detail, the central idea closely resembles Parsons and Shils' (PARSONS & SHILS 1965) concept of the unit act wherein interaction is defined by a mutuality of sanctions with regard to need gratification and deprivation, complementarity of expectations and the double contingency. The summary signed values represent the net payoff for each actor. By definition, each actor both receives and transmits sanctions.

The permutations of the binary values for the payoffs provide the basic relational states or environmental conditions which ET and NET analyse. As mentioned above, NET has largely been concerned with power arising in exchange.

+/+	Exchange
+/-	Coercion
-/-	Conflict

2. Define actors as rational decision making automata responding to environmental or external structural conditions. They comprise decision procedures, preferences and beliefs. Actors will always seek to satisfy their needs by acquiring resources through exchanges and will choose actions which maximise the return to their expected preference state of an exchange. They will avoid pay-offs which minimise their return. Failure to exchange is termed 'confrontation'.

3. Preferences are ordered sets of payoffs within a given structural state. As noted above, the bi-signed states form a triplet. An actor's preference order over this triplet is that actor's values. Mixed motive exchanges occur where preference orders differ across actors. Thus in a market exchange where the seller seeks the highest price and the buyer seeks the lowest cost, the set of preferred outcomes cannot satisfy both parties. Mixed motive exchanges are typical of weak power structures.

4. Actors are assumed to know their own preferences and to have beliefs about the preferences of their exchange partner. For simplicity of modelling, beliefs are assumed to be accurate.

The following axioms are derived from the postulates given.

1. In the simplest structure, the unitary dyad, A and B are the only exchange partners. Since there are no constraints or forces at work outside the dyad, no values or preferences are in play. Payoffs therefore are equal.[46]

2. In the idealised dyad, the structural power relation is equipower and the distribution of value through exchange is equal. The assumption of fixed value means the distribution of payoffs is

46 Like the inclined plane example we discussed earlier, this idealisation is the core conception and its use is comparative. More complex structures consisting of three of more partners are defined in terms of their conformity to or departure from it.

zero-sum. It follows that in mixed motive exchanges, an actor can only obtain an increase in payoff at the expense of the other. This is the definition of power.

3. Exchanges as singular. That is, actors exchange with one and only one partner at a time. Hence where there are competing partners for an exchange, since only one can be included in the exchange, other are excluded.[47]

4. In order to avoid being excluded, a potentially excluded actor will make better offers to potential exchange partners than an included one. This creates the possibility of asymmetry in payoffs.

On the basis of these axioms, NET proposes asymmetry creates the possibility of power difference where potentially excluded partners will receive less than equal returns.[48] Power structures are ordered (strong power -> weak power -> equipower) depending on the conditions of exclusion and inclusion in play. Power can be local, as in the dyad, or distal as in a chained set of exchanges. A can exercise power over C although A does not exchange with C but with B and B exchanges with C. This happens when C receives a payoff that is less than the arithmetic mean of the resources being transferred and A receives more than the arithmetic mean. If C's shortfall exactly matches A's gain, the A is in a strong power position relative to C.

From the above axioms, NET derives its central theorems.

1. An actor's predisposition to exchange is a function of two forces: the actor's *resistance* to exchange modulated by their *dependency* on the exchange. Resistance is defined as the ratio of what would be forgone by agreeing to exchange at a level below their preferred maximum to what they would gain above zero (the payoff at confrontation). This is expressed as the following function:

$$\mathrm{Ra} = \frac{(Pamax - Pa)}{Pa - Pacon}$$

Pamax is the maximal return to A; Pa is the offered return to A; and Pacon is the return to A at confrontation (by definition, 0).

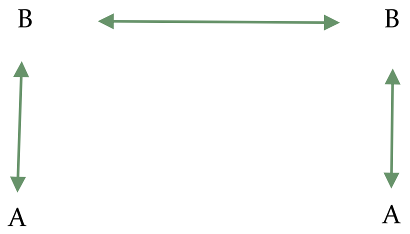
2. The point at which exchange occurs is where the resistances of both actors are equal, that is exchange occurs where Ra = Rb or

$$Ra = \frac{(Pamax - Pa)}{Pa - Pacon} = \frac{(Pbmax - Pb)}{Pb - Pbcon} = Rb$$

47 Latterly, NET has been exploring the consequences of relaxing this 'one at a time' constraint. See Willer & Emanuelson (2009)

48 Power in NET is, therefore, a structural property not a personal one. A is exercising power over B if B receives a lower payoff than might otherwise have been the case, even if (a) B is not aware of the shortfall or (b) accepts the shortfall willingly.

3. Dependency is a function of the likelihood of being excluded from an exchange and failing to exchange. Since by the scheme's definitions actors always prefer some payoff to no payoff, exchange will occur in situations of dependency even though payoffs are severely asymmetric.
4. In the dyad, A and B are the only partners and since by definition actors always seek to exchange the likelihood either will be excluded is zero. However, consider the following topology



Given all actors are predisposed to seek to exchange will all other available parties, each B will exchange with its A with a likelihood of 0.5 and with the other B with a likelihood of 0.5. Each A will seek to exchange its B with a likelihood of 1.0. Thus the joint probability Bs will exchange is 0.25. When the Bs exchange, the As are excluded. Thus *l*a, the likelihood an A will exchange is 0.75 and *l*b is 1.0.

Using likelihoods and resistances, we can predict different payoff structures for different power structures and topologies. For instance, the topology above is a weak power exchange. Both Bs have alternatives and so have some power. Since they have no alternatives, the A positions are in a weaker position. In a weak power arrangement, Pmax varies between the total resource value being exchanged and its mean value. Pcon varies between the mean value and 0. Applying the likelihoods of exclusion to Pmax and Pcon reflects the probability of inclusion and gives a qualitative estimate of likely pay-offs. To show how this works, assume the resource pool being exchanged is 24 units:

Thus, Pamax = $12(1+la)$ = 12 * 1.75 = 21 and Pacon = $12la$ = 12 / 0.75 = 9. This gives the resistance equation:

$$\frac{21-Pa}{Pa-9} = \frac{24-Pb}{Pb-12}.$$

Given Pa + Pb = 24, by iteration the solution is given as Pa = 10.5 and Pb = 13.5.[49] These pay-offs represent the compromise position adopted.

---

49 Note these numbers are not quantified measures. They represent more or less how much in a possible range of 0 to 24each party is likely to gain. The precision is or can be misleading.

Applications and extensions of the Theory

Although a number of historical settings (slavery in the Ancient World, capitalist labour relations, modern bureaucratic organisations) and invented scenarios (starting a company, visiting the doctor, moving furniture) are described using NET's terminology, these are really illustrative explications of the concepts; stories using the terminology rather than analyses of the specifics of particular social phenomena. The only actual structures which have been examined in detail are different graph topologies (stems, leaves and trees) constructed to display different power structures (that is, different constellations of exclusion and inclusion). The availability of computational tools has meant that these topologies can specified to involve much larger numbers of nodes than those previously examined in laboratory experiments. The extension of analysis to graphs with larger numbers of nodes has also been enabled by a more lightweight method of defining inclusion and exclusion based on path and neighbour counting.

The ability to manage large graphs has allowed NET to propose a method for decomposing large networks into sub-components based on the identification of "break points" where strong power positions are connected. (Willler et al 2012). Because strong power positions demand maximal payoff and the zero-sum postulate for pay-off distribution makes this impossible, it is hypothesised that strong power positions will not exchange and thus the edges delineating their relationships constitute "break points" within which weaker exchange partners comprise sub-networks. NET hopes this will provide a method for scaling up its analyses. By developing algorithms for the serial solving of resistance equations, attention is now being paid to networks where resources flow through a chain of partners. Finally, the constraints of the assumption of strict individualistic rationalism of conventional Exchange Theory have been recognised and attempts made to adapt the core theorems to enable the description of social structures using value orientations rather than payoffs (Willer et al 2013).

Discussion

The test of a formal theory like NET is generativity. From a small number of postulates, a social mechanism or mechanisms is constructed and theoretical descriptions of social phenomena recognised and studied within sociology can then be elaborated. Often this mechanism is quite surprising. A classic example of theorising this way is Thomas Schelling's (Schelling 1978) account of racial segregation. Using a few simple assumptions concerning preferred composition of neighbourhoods and available levels of housing mobility, Schelling showed that extreme forms of housing segregation based on race could result even when the original population is relatively mixed and everyone is 'highly tolerant'. Racial segregation in housing did not require members of communities to be 'bigots'. The patterning of segregation is highly recognisable and much studied by sociology. The mechanisms by which Schelling's theory produced it are, though, quite counter intuitive.

Measured against the Schelling standard, how does NET fare? Do its theorems produce generative descriptions of power in exchange that cause us to re-think current accounts of that phenomenon? Moreover, does it do so in ways that cause us to reflect on the mechanisms generating power differentials? Unfortunately, our answer is negative. The explanation of predisposition to exchange in terms of relative resistance is just a re-description in technical terminology of the kind of explanation most of us would give for the choices we make as economic or non-economic actors. Looking for "value for money" in purchases and being aware there is a point where "enough is enough" in the sharing of favours with friends, neighbours and acquaintances do not require us to conceive the social world as NET does and adds nothing to how we already think about our exchange relationships. The same is true for the theorem concerning the role of exclusion in facilitating power. If you have no options, then you have no options; and if you do, you do. NET's conception of power comes down, in the end, to an actor's capacity or willingness to choose to do otherwise, which in the circumstances the theory puts the actor, he or she cannot do.

If the derived descriptions are not informative, what about the mechanisms by which they are generated? Do they, like Schelling's, create recognisable social forms in ways we would not have expected? Once again the judgement must be negative. While there is no requirements that initial postulates be 'realistic' (whatever that means), the mechanisms by which it is proposed social arrangements are patterned and structured should be plausible as descriptions of how people carry on their lives. In Schelling's stories, we can readily see the ways individual ordinary people actually do behave. It is the ways that behaviour is aggregated that produces the phenomenon being theorised. As NET itself admits, its emphasis on a strict individual maximising rationality is directly at odds with how most of us experience social life. Introducing social values within a strictly rationalistic structure does not obviate the problem. (see Willer et al 2013).

The paucity of insight and innovation in NET derives from two flaws in its approach. First, the initial conceptualisation of power fixes it in terms of structural conditions. This is a deliberate disregarding of many aspects central to sociology's interest including the broadly institutional nature of power. What other sociological accounts conceive as a multivalent phenomenon is reduced to a single dimension. Most accounts of power see it realised through a variety of asymmetries using explicit and implicit control, influence and commitment etc. which are mediated and expressed in decentralised or institutional forms. Disregarding this required complexity is a perfectly proper approach to take if the resulting structures are generative enough to produce the complexity as the model and the phenomena it represents scale up. For many theories in sociology, this scaling is achieved by the diffusion of power relations throughout a social structure. Interpersonal and institutional exchanges may exhibit the consequences of power but are not where power lives, so to speak. Although NET triangulates its interest in power against the theories of Weber, Dahl, Lukes and others, it is hard to see how it could generate the scale of phenomena they are concerned with. Even with modern computational tools, dealing with chains of graphs for orders greater than 100 is a major challenge, particularly when the sequences of exchange iterate. If we want to build a cumulative sociology of power in which the key contributions of the theory of alienation, the theory of bureaucracy, structuration theory and so on can be closely coupled within Exchange Theory, then NET has to provide ways to translate conceptualisations central to those theories into its formalism.[50]

Second, the theory fails to take more than minimal advantage of the formalism it adopts (Graph Theory). Hardly any use of made of the range of the mathematical concepts, axioms and theorems which have been developed there. Instead, simple graphical structures are used to depict (not analyse) specific structures. The result is that the formalism operates as a loose not formal analogy. The same is true of the key resistance functions which are, in fact, functionally expressed similarly informal analogies of common sense understandings.

The Proponents of NET make much of the novelty of the approach and the analytic power it offers sociology. And yet there is nothing intrinsically innovative about seeing power (or influence and control) exercised over chains of relationships, nor is there anything intrinsically new in conceiving actors as operating minimaxing strategies. Certainly, the formalisms add nothing to the sociological specification of what these strategies might amount to. No new mechanisms are adduced and nothing is added concerning what actually goes on in the exercise of power and a great deal of subtlety and sociological insight is lost concerning how alliances are formed, the ways trade-offs are interpreted and re-interpreted and how obligation, prestation, kinship and similar relationships operate within exchange systems. NET only tells us that *somehow* exchange relations take place in such a way to produce the payoffs that they do. This means it is altogether empty of insight as to what exchange relations are actually like.

Of course, the shallowness of NET is not a demonstration of the futility of qualitative formal theorising, only of the challenges theorising in this way faces and the requirements we might have of it. It is not enough to assert "Treat social relationships as exchange" or games or networks or whatever and then proceed to describe social phenomena as if they conformed to that assertion, whether that description be formal or informal. Rather what is required is the analysis of the mechanisms made available *if* the assertion is adopted which shows how they generate the phenomena of social life which sociology addresses and how those generative mechanisms throw light on how society might operate.

50 Interestingly recent attempts to show how scaling might be achieved work in precisely the opposite way. What is proposed are ways of de-composing (or 'scaling down') large scale networks into smaller ones not multiplying smaller units up (see WILLER ET AL 2012).

VII. CONCLUSION

Formalisation is conventionally seen as the use of the tools of logic or mathematics to provide rigorous reasoning. It is a means to theorising. As such, it depends upon the formal regimentation of a set of theoretical propositions. However, first, the propositions must be capable of being regimented in the ways stipulated. And thereby hangs the tale for sociology. The current constitution of much, if not most, sociological theory not only resists formalisation, it positively militates against it. What it wants to be is precisely what formal theory is not. Given this at best misalignment and at worst contradiction, the pursuit of formalisation poses sociology a stark choice. If this is what formalisation is, either sociology can continue theorising in the ways it currently does, with all the formal imperfections and all the informal advantages it currently has, or it can transmogrify itself into a narrowly framed discipline akin to axiomatic economics or experimental psychology.[51] This will definitely change the forms of theory which sociology compiles and allow mathematical formalisation. It will also change the forms of data it collects and the way it collects them. Theory and method are inextricably linked and the changes will be far reaching. For some, this price might be worth paying. Given the discipline's reluctance to conform to the blandishments of the advocates of formalisation, we suspect that they are very few and probably reducing in number. The momentum seems to be all the other way.

A second set of considerations would also caution us to step carefully here. These have to do with what we might broadly call the 'formalisation-readiness' of sociological theory. As we have seen, even if it was not against the grain of most sociological theorising, the current state of that theory would make the process of formalisation at best pointless and at worst obfuscating. In addition, formalisation would not, by itself, actually contribute to the realisation of sociology's overarching ambition, namely to provide accounts of the social world which map with increasing precision on to social reality. What matters is the degree of precision in the mapping. The test of such precision was and is empirical. Empirical demonstration plays no part in the process of formalisation since emphasis is exclusively placed on the structure within which theory is represented. The tests applied have to do with truth values and valid deduction. In the natural sciences, truth values and empirical demonstration are connected by (formal) theories of measurement, key to which are assumptions about the and stability and independence of effects along different causal paths well as the SUTVA principle. Given isomorphic models and formally constructed measurement systems, formalisation offers buttressing support for the credibility of empirical statements. Because of the openness and fragility of the models constructed and

---

51 As our phrasing might intimate, we do not think this is all that formalisation might be. In our view, the description of the turn by turn operation of conversation given in SACKS, SCHEGLOFF AND JEFFERSON (1974) is a formal description. We would even be prepared to call it a formal theory. That Conversation Analysis has failed to take this effort forward (and where it has not stalled, has actually gone backwards) is a reflection of the state of Conversational Analysis not the rigour, innovativeness and generativity of the model presented.

the lack of formalised measurement theories, at present in sociology there is no way to connect formalisation of theory and empirical demonstration. If we want buttressing evidence for the credibility of our theoretical claims, simply trying to formalise our non-formal theoretical models will not help. As a result, formalisation alone will not contribute to the aspiration for sociology to become more like the natural sciences. *At this point in the discipline's history*, the pursuit of this kind of formalisation is pointless, something that those arguing for formalisation actually seem to acknowledge through the illustrative nature of their efforts — this is what sociology would look like if only it were very different in many other respects beyond just those introduced by the formalising exercise.

In addition to the fit between the requirements of formalisation and what sociology can currently offer, there are considerations concerning the consequences of formalisation itself. Initially formalisation will make little or no difference to the character of sociology as it is practised. But as the formal discipline evolves, its character will undoubtedly change. Formalising what are already complex theories tends to raise the order of complexity. As can be seen from the relatively trivial examples we have used, these complexities ramify exponentially once we move beyond the most basic cases. Not only will this make sociological theory hard to read, it will make it hard to survey and even harder to summarise. The formal mathematical theories used in physics often run to 100s if not 1000s of pages. At this scale, it is a consummate and hard won skill to be able to grasp all the components, let alone survey the totality to understand its level of coherence and integration. Most current sociological theory might be formally feeble from a mathematical and logical point of view but at least it is transparent and surveyable.

Our fourth observation is more a summary comment on the cases we have studied than a conclusion. What is remarkable about the pursuit of formalisation is how quickly the central sociological concerns are moved into the background in favour of issues to do with the structures of particular formalisms or the operation of logic. Somehow the sociology (which is, we presume, what the formalisation exercise is all about) is left to take care of itself, something which it has signally failed to do. In the end, any connection between the communicative transparency and rigour which formalisation might bring to sociology and the sought after traction on sociological problems and issues is hardly, if ever, made.

We don't want to be a wholly negative though. We should remember formalisation is in the service of proof, and what counts as a proof is changing or at least under discussion. Recently theorems in mathematics have been 'proved' by the brute force use of high power computing (see Krantz 2011). True not everyone accepts these proofs as pukka, but the discussion is on the table. What this might imply is that if we can find some mathematicians like John Kemeny who are sufficiently interested in our sociological problems to design bespoke mathematics for us (rather than as we noted earlier, sociologists with an attraction to mathematics) perhaps they will also be able to invent new forms of proof to go with them; forms of proof which match what we want to do and what we can do.

In the end, the conclusion we draw from this discussion is much the same as the conclusions we have drawn in our other discussions of the possibility of a scientific sociology. It is easy for influential bodies and senior members of the profession to say sociology should be quantitative, formally mathematical, or adopt the methods of science. Showing how that can be done in a way which is more than merely notional whilst preserving the values which motive the discipline is something else entirely. And yet is it precisely such demonstration which is required. Rather than exhortations, vague programmatics, sketch ideas and promissory notes, what is needed is for those who advocate a truly scientific sociology to pick up the task of working through the range of issues and challenges the proposal faces in order to show the rest of us how it can be done.

In failing to provide such a demonstration, we wonder if the advocates of formalisation and the adoption of one or other scientific model can actually see the scale of the challenges. Clearly it involves more than just a small matter of notational translation and a little light axiomatization.[52] Those who think that way signally fail to appreciate the basis on which the natural and mathematical sciences have made the progress they have. As Herbert Hochberg (Hochberg 1958) dryly noted long ago, physics has the panoply of powerful theories it has because it has done the painstaking work needed to show that its concepts can be fully interpreted within an array of interrelated mathematical formalisms and anyone who thinks that social science can achieve the same result without undertaking the same painstaking work would be foolish indeed. But there again, a strategy of taking one small but secured step at a time in pursuit of a long term disciplinary vision has long been anathema in sociology.

[52] We wonder how far this perception is associated with the desire of professional and other bodies to produce research and researchers of use in the formation of bureaucratic policy and so only required to do basic data collection and analysis.

REFERENCES

Anderson, R & Sharrock W.	2013	*Analytic Sociology*. http://www.sharrockandanderson.co.uk/the-archive/1990-present/post-2010/
Anderson, R & Sharrock W.	2014	*Empirical Philosophy*. http://www.sharrockandanderson.co.uk/the-archive/1990-present/post-2010/
Anderson, R & Sharrock W.	2015	*Visualisation*. http://www.sharrockandanderson.co.uk/the-archive/1990-present/post-2010/
Barnes, J & Harary,F.	1983	Graph Theory in Network Analysis. *Social Networks*, vol 5, pp 235 - 244.
Berkman, L., Kawachi, I. & Glymour, M.	2014	*Social Epidemiology*. Oxford. Oxford University press.
Blalock, H.	1994	Why Have We failed to Systemize Reality's Complexities? In Hage *Op Cit.*, pp 227 - 242
Blalock, H. (ed)	1985	*Causal Models in the Social Sciences*. New York. Aldine.
Blumer, H.	1956	Sociological Analysis and the "Variable". *American Sociological Review*, vol 21, pp 683 - 690
Brodbeck, M.	1958	Models, Meaning and Theories in L. Gross (ed.) *Symposium of Sociological Theory*, New York. Harper Row, pp 373-406
Cartwright, N.	2007a	*Hunting Causes and Using Them*. Cambridge. Cambridge University Press
Cartwright, N.	2007b	Are RCTs the Gold Standard? *BioSocieties*, vol. 2, pp 11 - 20.
Cicourel, A.	1964	*Method and Measurement*. New York. Free Press
Duncan, O.	1975	*Introduction to Structural Equation Models*. New York, Academic Press.
Davis, J.	1994	What's Wrong with Sociology? *Sociological Forum*, vol. 9, no 2, pp 179 - 197
Easley, D. & Kleinberg,J.	2010	*Networks, Crowds and Markets*. Cambridge. Cambridge University Press.
Elwert, F & Winship, C.	2014	Endogenous Selection Bias: the problem of conditioning on a collider variable. *Annual Review of Sociology,* vol. 40, pp31-53.
Elwert, F.	2013	Graphical Causal Models in S. Morgan (ed.) *Handbook of Causal Analysis*. Dordrecht. Springer. Pp245 - 273
EPSRC	2010	*International Benchmarking Review of UK Sociology*. Swindon. EPSRC
Freese, L.	1980	Formal Theorizing. *Annual Review of Sociology*, vol. 6, pp 187 - 212.
Furfey, P.	1954	The Formalisation of Sociology. *American Sociological Review*, vol. 19 no 5, pp 525 - 528

Garfinkel, H	1967	*Studies in Ethnomethodology*, Englewood Cliffs, Prentice-Hall.
Greenland, S., Pearl, J. and Robins, J.	1999	Causal Diagrams for Epidemiological Research. *Epidemiology*, vol. 10, no 1, pp 37 - 48.
Hage, J. (ed.)	1994	*Formal Theory in Sociology*. New York. SUNY Press.
Hammersley		*BSA Annual Conference Programme and Abstract*
Harary, F.	1969	*Graph Theory*. Reading, Mass. Addison-Wesley.
Heylighen, F.	1999	Advantages and Limitations of Formal Expression. *Foundations of Science*, vol. 4, no 1, pp 25 - 56
Hochberg, H.	1958	Axiomatic Systems, Formalization and Scientific Theories in L. Gross (ed.) *Symposium of Sociological Theory*, New York. Harper Row, pp 407-436
Holland, P.W.	2001	Counterfactuals in Social Science. Neil Smelser & Paul, B. Baltes (eds.) *International Encyclopaedia of the Social and Behavioral Sciences*. Amsterdam. Elsevier Pp 1550 - 1554
Hutchinson, P., Read, R. & Sharrock, W.	2008	*There's No Such Thing as Social Science*. Aldershot. Ashgate.
Kamps, J.	2000	*A Logical Approach to Computation Theory Building with Applications to Sociology*. Amsterdam. Institute for Language, Logic and Computation
Kemeny, J.	1959	Mathematics without Numbers. *Daedalus*, vol. 88, no 4 pp 577 - 591
Knight, C. & Winship, C.	2013	The Causal Implications of Mechanistic Thinking: Identification Using Directed Acyclic Graphs (DAGs). In: Morgan, S. *Op Cit.*, pp 275-301
Krantz, S.	2011	*The Proof is in the Pudding*. Dordrecht, Springer
Leiberson, S.	1985	*Making it Count*. Oakland. University of California Press.
Leiberson, S. and Lynn, F.	2002	Barking Up The Wrong Branch. *Annual Review of Sociology*, vol. 28, no 28, pp 1 – 19.
Lewis, D	1973a	Causation. *Journal of Philosophy*, vol. 70, no 17, pp 556 -567
Lewis, D	1973b	*Counterfactuals*. Oxford. Blackwell
Lleras, C	2005	Path Analysis. *Encyclopaedia of Social Measurement*. Vol 3 Amsterdam. Elsevier. Pp 25 - 30
Mirowski, P.	2002	*Machine Dreams*. Cambridge. Cambridge University Press.
Morgan, M & Morrison, M (eds.)	1999	*Models as Mediators*. Cambridge. Cambridge University Press.
Morgan, S. & Winship, C.	2007	*Counterfactuals and Causal Inference*. Cambridge, Cambridge University Press.
Morgan, S. & Winship, C.	2012	Bringing Context and Variability Back into Causal Analysis in H. Kincaid (ed.) *Oxford Handbook of the Philosophy of Social Science*. Oxford. Oxford University Press. Pp 319 - 354
Morgan, S. (ed)	2013	*Handbook of Causal Analysis*. Dordrecht. Springer.
Mueller, E.	2006	*Commonsense Reasoning*. New York. Morgan Kaufmann

Neyman, J.	1935	Statistical problems in Agricultural Experimentation. *Journal of the Royal Statistical Society*, Series B no 2, pp 107 - 180
Parsons, T., & Shils, E. (eds.)	1965	*Towards a general Theory of Action*. New York. Harper-Row.
Pearl, J.	2010	The Foundations of Causal Inference. *Sociological Methodology*, vol 40. pp 74 - 149
Pearl, J.	2009a	*Causality*. 2nd Edition. Cambridge. Cambridge University Press.
Pearl, J.	2009b	Causal Inference in Statistics. *Statistical Surveys*, vol 3, pp 96 - 146
Quine, W. V. O.	1969	'Ontological Relativity' in *Ontological Relativity and Other Essays*. New York, Columbia University Press, pp 26 – 58.
Rubin, D.	2005	Causal Inference Using Potential Outcomes. *Design, Modelling, Decisions*. Journal of the American Statistical Association, vol. 100 no 469 pp 322 - 331
Russo, F, Wunsch, G & Mouchart, M.	2011	Inferring Causality in Counterfactuals in Observational Studies. *Bulletin of Sociological Methodology*, vol. 111, no 1, pp 43-64
Russo, F. (ed.)	2009	*Causality and Causal Modelling in the Social Sciences*. Amsterdam, Elsevier.
Sacks, H., Schegloff, E. and Jefferson, G.	1974	'A simplest systematics for the organisation of turn taking in conversation.' *Language*, vol 50, pp 696 – 735.
Schonemann, P	1994	Measurement: The reasonable ineffectiveness of mathematics in the social sciences in Borg, I & Mohler, P. Trends, *Perspectives in Empirical Social Research*. Berlin, De Gruyter. Pp 149 - 160
Schutz, A.	1962	Common-Sense and Scientific Interpretation of Human Action. *Collected Papers*, (M. Natanson ed.) The Hague. Martinis Nijhoff. Vol 1. Pp 3 – 47.
Shalizi, C. & Thomas, A.	2011	Homophily and contagion are generically confounded in observational social network studies. *Sociological Methods and Research*, 40, 211–239.
Sharkey, P & Elwert, F.	2011	The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. *American Journal of Sociology,* vol. 116, no. 6, pp. 1934 – 81.
Schelling, T.	1978	*Micromotives and Microbehaviour*. New York, W.W. Norton & Company.
Smith, H.	2014	Effects of Causes and Causes of Effects: some remarks from the sociological side. *Sociological Methods and Research*, vol. 43, pp 406 - 415.
Thompson, J.	1967	*Organisations in Action*. New York. McGraw-Hill
Walker, H, Thye, S., Simpson, B., Lovagalia, M., and Willer, D.	2000	Network Exchange Theory: Recent Developments and New Directions. *Social Psychological Quarterly*, vol. 63, no 4, pp 324 - 337

Willer, D.	1996	The prominence of Formal Theory in Sociology. *Social Forces*, vol. 11, no 2, pp 319 - 331
Willer, D.	2003	Power-at-a-a-Distance. *Social Forces*, vol 18 no 4 , pp1295 - 1334
Willer, D. (ed.)	1999	*Network Exchange Theory*. West Port, CT. Praeger
Willer, D. and Emanuelson, P.	2006	Elementary Theory in P. J. Burke, *Contemporary Social Psychological Theories.* Stanford. Stanford University Press, pp 217 – 243.
Willer, D. and Emanuelson, P.	2008	Testing Ten Theories. *Journal of Mathematical Sociology*, vol. 32, no 3, pp 165 – 203.
Willer, D. and Emanuelson, P.	2009	One shot exchange networks and the shadow of the future. *Social Networks*, vol. 31, pp 147 – 154.
Willer, D. van Assen, M, and Emanuelson, P.	2012	Analysing Large Scale Exchange Networks. *Social Networks*, vol. 34, no 2, pp 171 – 180.
Willer, D., Gladstone, E., and Berigan, N.	2013	Social Values and Social Structure. *The Journal of Mathematical Sociology*, vol. 37, no 2, pp 113 – 130.