4

Two Puzzles about Mathematical Sociology

Introduction

The first of our puzzles is a technical issue in the application of mathematics in the sciences. The second is more a professional matter. We have some thoughts about the technical issue and these will be the focus of this essay. We don't know what to think about the professional matter and so, rather than venturing opinions, we'll just flag it up.

The puzzles are these. A central thread in much of the debate in the Philosophy of Science concerns the realism of mathematical structures. It is central because almost everyone agrees the application of Mathematics in the natural sciences is both a reason for their undoubted success and one of their distinguishing features. The trouble is, and this is the issue, no-one knows how to justify the claim that the mathematics used captures aspects of 'the reality' of whatever is being investigated. That is to say, no-one has so far formulated a good *philosophical* argument for how or why purely formal abstractions are able to represent material goings-on.[1]

Though perhaps not quite the scandal the failure of Philosophy to answer Hume was to Kant, nonetheless the situation has left a lot of philosophers more than a little uncomfortable. The results of the applied sciences seem to show the mathematics must work somehow. After all, to cite

[1] We sometimes talk of material goings on as substantive and mean by this spatio-temporal. This is fine for the natural sciences but poses problems for the social and psychological. For example, in what way are the institutions of Government or Justice spatio-temporal? As for thoughts and the processes of thinking.....

the favourite examples, science has put people on the moon, developed vaccines using gene editing and is building quantum computers. Since the probabilities of the mathematics deployed just happening to fit the uses to which they are put are so astronomically small, invoking them as justification would make the successes akin to a cosmic miracle. Alternatively, no-one wants to rely on luck for an argument. These days, somewhat reluctantly, the consensus seems to be to default to what is called an inference to the best explanation. The mathematics just works, even if we don't know why. [2]

The professional aspect of this is something of a dog that didn't bark. We would not claim to be professional philosophers of social science. Neither are we members of the branch of Sociology called Social Studies of Science and Technology. But we are interested outsiders and do follow the debates in both domains reasonably closely. The thing is, in all the discussion over the foundations of modes of Sociology, their metaphysics and ontologies, scientific status, the nature of their scientific practise and so on, we have not come across any consideration of the philosophical puzzle about the applicability of mathematics and what that might mean for the practise of Sociology. We have seen lots of other—usually critical—discussions, but not that one. Two of the implications, or so it seems to us, are pretty obvious. The first concerns the use of mathematical methods and structures to organise sociological descriptions. The second has to do with how in practice the mathematics is put to work in Sociology as well as in the (paragon) sciences. The former is, of course, just the extension to Sociology of the philosophical puzzle. The latter points to a whole research domain which, paraphrasing Andrew Pickering [1995], we could summarise as 'the mangling of mathematics'.[3]

While there are no generally agreed positive philosophical arguments to resolve the mystery, there are some technical mathematical ones. We will use one as our point of departure. Employing it, we will offer a sociological account of the philosophical puzzle by suggesting one way of solving the puzzle could turn on sociality. That is to say, we will sociologically construe the mathematical practise the solution represents and propose therein lies a way out of the quandary. We show how this might be done by looking at a fairly routine example of the application of mathematics to social phenomena. Our argument goes like this. When applied to the social, at

---

[2] The list of the philosophers who sign up for this inference to the best explanation is long and the status of many of them impressive. These are not people who don't know a good bone to gnaw on when they see one. And yet they don't. Perhaps that's why their sociological counterparts seem to be asleep at the switch.

[3] Just to be clear. Eric Livingston has already led the way here. See [Livingston, 1986; 1999]

least, what the philosophical puzzle references is the possibility of an abstraction gap between the mathematical formulation of a sociological phenomenon and the phenomenology or lived experience of some part of social life for which that phenomenon is a sociological reconstruction. When looked at sociologically, the task or work of bridging the abstraction gap is part and parcel of doing scientific or sociological description in a professional and competent manner. This competence, we will assert, involves double fitting the mathematics and the phenomenology to one another.[4]

Section 1. Mapping Mathematical Structures

Of course, the realism of Mathematics is a sub-issue in the general concern with what doing mathematics actually involves and what its foundations are. There are numerous offerings here, some of which go back to the Pre-Socratics. One relatively recent one, Structuralism, is the version we want to focus on. At its simplest (and this is very, very simple!) structures are classes of mathematical sets. Mathematics is about is the "mapping" of such sets. Here is how Stewart Shapiro, one of its modern proponents, explains the idea.

> Define a *system* to be a collection of objects.....Define a *pattern* or *structure* to be the abstract form of the system highlighting the interrelationships among them, and ignoring any features of them that do not affect how they relate to other objects in the system [Shapiro, 2000 p.259 italics in original].

Structures, then, are abstractions over configurations of the relational positions of objects. Any actual configuration can be patterned in a number of ways. So, the coloured stripes on the cover of a notebook may be seen as an ordered sequence of green, red, black, white and yellow lines or as blocks of cyan, magenta, yellow and black printed dots or as an open representation of emotional states. For the mathematician, it does not matter what the *relata* of the positioning are (ordered stripes, clusters of pixels, emotional responses). The abstract relational structures are what mathematicians are concerned with. And what they want to do is prove theorems about them.

A mapping defines an isomorphic relationship between structures: the objects of one structure can be equally successfully represented by another. To take a straightforward example. The Integer Numbers ($\mathbb{Z}$) 0,1,2,3,4,5 can be mapped into the Rational Numbers ($\mathbb{Q}$), ½, ¼, ¾ etc, by forming ratios of non-zero integers. $\mathbb{Z}$ and $\mathbb{Q}$ are different mathematical structures with a

[4] Here lies the implication of this line of thinking for the Sociology of Knowledge in its broadest sense. The double fitting of descriptions might just be a general enough practice to resolve rafts of problems in social epistemology which otherwise seem intractable. See [Baldamus 1971].

mapping. What ensures the success of a mapping is that it is 'structure preserving'. This notion is extended to applied Mathematics by proposing successful mathematical structures preserve the inherent structure of the domain of application. The way this is this done is by finding a minimal set of features or properties which can be used to characterise some physical object or process (see Jessica Wilson [Wilson 2010]). Different colours of the rainbow, for instance, might be reduced to wavelength frequencies. The heat of a kettle and the journey of a car might be reduced to point masses in motion. The purpose of characterisation is to ensure nothing of relevance to the problem's analytic framework is left out and nothing irrelevant included. Effective characterisation is what delivers structure preservation.[5]

Let's take a familiar example from the social sciences and sketch how the idea might apply there. Suppose we are interested in why Governments are reluctant to tax high earners at greater rates than low earners. The answer usually given (and we are not engaging in a debate over this) turns on the economic object 'the consumption function' and the relationships between its major components, autonomous consumption and the marginal propensity to consume. Autonomous consumption is consumption necessary to maintain basic social existence. The marginal propensity to consume is the increase or decrease in consumption as income rises beyond the level needed to maintain autonomous consumption. Economics formulates the consumption function and its components as mathematical objects such as numerical levels, rates and proportions. For families on low income, autonomous consumption requires all or almost all disposable income. As income rises, the proportion spent on autonomous consumption reduces and the residue can be allocated to other sets of goods and services. Eventually (and this turns on the psychology of utility functions—another mathematised concept—but we won't get into that), the very well off begin to run out of things they wish or need to spend money on and so save. Savings are a large element of the proportion of National Income available for investment. Investment is held by most Governments to be the most important driver of economic growth and stability. Hence, they are reluctant to reduce the 'willingness to save' of high earners by reducing their disposable income through taxation.

What is happening with the adoption of the consumption function as the rationale for economic policy is an assumption about a commonality between the structure of people's social and economic choices and that of a mathematical object, the consumption function. Income,

---

[5] Effective is not the same as exhaustive or perfect. Characterisation of physical processes by reducing them to other physical processes does not provide for an object's category hopping from being spatio-temporal to an abstract one. The end-to-end closure problem remains and hence the mystery.

consumption and saving are measured by translating goods and services accessed and consumed into monetary values. These are expressed in the Real Number system and the consumption function is a function applied to the result. By means of the consumption function, reasoning about how people deploy their incomes within the social system of economic daily life is mapped into reasoning about Real Numbers. The claim is that the structure of the function maps onto and preserves the structure of the spending. But what, exactly, is involved in doing this? That is the mathematico-philosophical question.

In a recent discussion of the debate, James Nguen and Roman Frigg [2022] suggest examples like ours raise three distinct questions.

a. The first has the general form of the philosophical puzzle we asked at the start. How do mathematical structures (a mathematical description) of any kind map? What general conditions are required for them to do so? This is not about prediction. A function will map if it generates mathematically sound results no matter whether they turn out to correlate with the way the social and natural worlds work. Any mapping will have to satisfy appropriate mathematical criteria.
b. How does a particular mapping work? What particular conditions are required for it to work? Is the structure the right one? What is the right form of the function? Are there limitations to its use, and so on?
c. Then there is the fundamental domain question. What enables facts about a mathematical structure expressed as the consumption function (say) to *explain* (if it does) facts about people's spending?

Nguen and Frigg—they are philosophers, after all—are only interested in the first, the philosophical, question. We are interested in the second but will use Nguen and Frigg's analysis as a platform for setting up our interest. The 'explanatory' can with all of its writhing contents will stay closed.

Nguen and Frigg's conclusion is the mapping account doesn't answer the philosophical puzzle. This is because the target of the mapping is usually taken to be a data model of the system of relationships and data models are themselves already mathematical structures. They are mathematical models derived by applying measurement functions on physical and social goings on. So, while the mapping is a perfectly proper mathematical description, we are still left with the

original problem, except now it is about data models. How do abstract mathematical sets in the form of data models map onto non-mathematical substantives?[6]

Nguen and Frigg's diagnosis of why the mapping account in this form doesn't work is straightforward. It just doesn't go far enough. They propose an amendment involving the adoption of what they call their "extensional abstraction account" [2022 p. S5953]. When investigators undertake a study, they have to decide which features of 'the world' they wish to describe and how. In sociological jargon, this is the resolution of the problematic possibilities of description. Whilst we might think it is quite natural to see the set of transactions making up the economy as direct and indirect exchanges of valued goods and services facilitated by media of exchange, there are certainly other ways they can be conceived economically (e.g., circulation of surplus labour value). The standard model is just that, a standard model. It is not the only model. The model selects certain orders of objects and the relations they stand in. It sets aside other orders of objects and their relationships. The description given is an instantiation of that selection.

Howsoever we choose to describe it, the description of actual economic activity has to be processed by a step by step climbing of the "abstraction ladder". At every move from riser to riser, more of the directly domain-relevant features of the description are stripped out. Eventually, nothing pertaining to the domain is left. All we have are objects coded in a mathematical structure.[7] But now the question becomes how to move back to the domain when we have the results of the mathematical operations on the structure which we generated. How do we step down the ladder again? This can only be done if we recognise our results depend on the structure generating original (selective) description. The mathematics of the consumption function does not capture essential reality, some uniquely identifying structure of economic life. It is a rendering of economics in a chosen modality. In earlier work, Frigg and Nguen [2019] called such renderings "keyed descriptions".

Nguen and Frigg's solution to the puzzle of the applicability of any mathematics rests on assumptions about the relative openness of selectivity of choice and its institutionalisation in a

---

[6] This is, of course, simply looking at the relationship in terms of the ability of the mathematics used to respond to the constraints of the physics being described. Noah Stemeroff recently pointed out the constraints that the mathematics places on the descriptions the physics can give are just as important for the viability of the mapping. See [Stemeroff 2021].

[7] This stepwise abstraction is a very good explication of what Woodward and Bogen [Woodward and Bogen 1988] were pointing to in their discussion of scientific practise as the transformation of physical data into theoretical phenomena.

professional practise. The latter can be rendered in a standard sociological form. For the sociologist, the patterns of empirical description are institutionalised when they are normatively regulated. They have a normative social character which is open to analysis under an appropriately construed sociology.

Applying the problem as characterised by Nguen and Frigg to the practise of Sociology causes no difficulties for sociological investigators. They just don't view it as a mathetico-philosophical question and so don't talk about it in those terms. For them, it is a practical question; one of operationalisation through coding. Somehow, the collection of reports, records, responses, notes and other impedimenta capturing the investigative experience of any study has to be processed and rendered in some appropriate notation to make it amenable to a chosen mode of analysis. In many cases, just as with the consumption function, this notation is derived from the Real Number system. What the coding produces is the data model to which functions on the Real Numbers can be applied. We used the term "somehow" just now. We did not mean "just anyhow". What the sociological rendering picks out are the disciplinary expectations which normatively regulate research practise and reporting. What these are and how they work are what sociological investigators discuss. Except mostly they don't do it explicitly in the reports they write. Instead, what we have called "the plausibility structure" of reporting (its demonstrable conformity to professional expectations of competent practice) is carried in the reporting itself. For those who know how to read for it, the reported reconstruction provides an "account" of its own production. This is what we meant by suggesting there is a sociological solution to the philosophical problem. Displayed competent practice instantiated in managing the coding problem provides a "for all practical purposes" bridge across the abstraction gap. In the next section, we illustrate one way this can be made visible.[8]

Section 2. Attitudes to Governmental Intervention

In a recent paper, Andersen, Curtis and Brym [2021] model how support for governmental social and economic intervention varies with levels of societal prosperity and income inequality. They

[8] The "for all practical purposes" qualification here is important. The solutions used are not philosophically inspired nor philosophically justified arguments. They do not answer the question which worries philosophers. What they do is manage the problem by coordinating a practical solution as the recognisable accomplishment of plausible coding. To find a practicable means of avoiding the philosophical impasse, sociologists shape both the data and the structure to get them to have a 'good enough' fit to one another and they do so in recognisably professionally competent ways. Just for completeness sake and to prevent sociologists from incurring too much angst, we should point out that John Bell, one of the foremost theoreticians of Quantum Theory, complained much, if not all, of standard text book accounts of Quantum analysis is of this loosely connected, for all practical purposes kind.

make extensive use of official or semi-official large-scale data sets. The base data used for the study have been collated from cross national surveys of 66 nations carried out by a number of Governments, supra-national Agencies and NGOs. The work required to mould this data to the problem they set themselves is what we are interested in.

Problem Statement

The core of the report consists in the consideration of a number of alternative regression models for the distributions of attitudes. The text defines the problem which the investigators wish to address as the "contradictory" nature of the extant literature on attitude formation towards governmental socio-economic intervention. This contradictoriness arises because of an insufficiently precise focus on the subjective drivers of attitudinal formation. Clarifying these drivers will, or so Andersen et al. propose, facilitate resolution of the contradictions.

Registration

The sought-after clarification is achieved by describing the three-way relationships between attitudes, income inequality and prosperity. Under the current, contradictory approaches

> ....we expect support for government intervention to be high if inequality is high, regardless of the level of prosperity. We also expect support to be high if prosperity is low, regardless of the level of income inequality. On the contrary, support for government intervention should be lowest under high prosperity and low inequality [Andersen et al. 2021 p. 1351]

This expected pattern is illustrated by an associated representation.

TABLE 1 Expected public support for government intervention by level of economic prosperity and income inequality

Income inequality		Economic prosperity: *Low*	Economic prosperity: *High*
	High	High	High
	Low	High	Low

Figure 1 From Andersen et al p.1351

The intuition is an obvious one. People's attitudes to Governmental socio-economic policies are driven by self-interest and reflect two socio-economic factors: income inequality and economic prosperity. The general level of support can be defined by positions on the nominal co-ordinates of the dimensions. This is the initial conceptualisation of the problem. Attitudes are locations in a Cartesian space of socio-economic dimensions.

Characterisation

'Box diagrams' of this kind are widely used in Sociology to simplify complex relationships and frame empirical and theoretical investigations. Perhaps the most famous are Merton's "Theory of Deviance", Parsons' "GAIL" structure of Functional Imperatives and the 'Prisoner's Choice Paradigm' in Game Theory. Andersen et al.'s use of this heuristic allows the reader to visualise how to step through the intuitive pattern they have hypothesised. In that sense, while *as a statement of the problem space* the illustration adds nothing to the text, it greatly enhances the comprehensibility of interaction in that problem space. Analytically, the text and the table accomplish precisely the same tasks. All the table does is 'concretise' the pattern by stripping out ancillary detail.

The text/picture combination achieves several things:

1. Isolation of the problem space;
2. Reduction of the problem statement;
3. Definition of a first order putative solution to that problem.

We'll take each of these in turn.

1. Attitudes to Government policies have all sorts of social and social psychological sources and supports. Much work has been done trying to articulate the processes by which public and personal opinion are shaped by social context at a personal and societal level. The approach Andersen et al. summarises simply sets that work aside. From the welter of contributing forces, we have a condensate of just two societal levels of income inequality and economic prosperity. These two massively simplify the degrees of variability in the problem space. We have just two dimensions and two levels of measurement.
2. The problem statement is encapsulated in the patterned contents of the boxes: the high/low mappings. This fixes the report's initial explanadum. How are the attitudinal outcomes in the boxes related to the social conditions?
3. The initial explanans is a mechanistic one. Relative position on the dimensions determines the attitudinal outcomes.

To find a place to insert their clarification, Andersen et al modify the explanans. The degree of support for intervention varies with *personal perception* of income inequality. In societies with high inequality and high prosperity, support for intervention is greatest among that segment of society which believe they are relatively poorly off. In other words, once you factor in personal

perceptions of one's own position in the income structure, the contradictions are resolved. In this modified explanation, there are three input social factors in play: levels of income inequality and societal prosperity as before and interpretations of 'relative deprivation'. Now we have three dimensions.

Notation

The influence of the input factors on attitudes is the sociological phenomenon they wish to describe. They model this as a series of regression functions. These models will provide instructions for how to read and what to read into the pattern set out in Figure 1. Regressions are partitions of the 'contributions' individual co-variates make to the overall variance of a variable. The clarificatory solution Andersen et al offer is arrived at by applying a regression function on a data model representing perceived and actual income inequality and economic prosperity together with policy attitudes held by the surveyed individuals. The structure of measures of co-variance using a notation in $\mathbb{R}$ is the data model Andersen et al. need to create to allow the regression functions to operate. To get that, they have to operationalise and abstract over the perceptual data.

Problem Specification

System Laws

Go back to the original problem statement. The propositions are couched as generalised relationships among theorised socio-economic phenomena. They reference individuals' pre-theoretical phenomenal social experience. Economic prosperity refers to the array of goods and services available for purchase. Income refers to an individual's capability to access those goods and services. Attitude to government intervention refers an individual's view of the role of Government action to redress socio-economic inequality. Finally, a sense of one's own position in the overall distribution of income is quite obviously a personal judgement.

In the problem statement, the two socio-economic drivers determine the sense individuals have of their social experience. This relationship is expressed mathematically as an Ordinary Least Squares (OLS) regression equation in which the value of dependent variable is a linear function of the independent variables. In their presentation of their statistical models, Andersen et al. detail how they teased apart the data by using various control mechanisms to produce different models for mixed effects and country specific effects across the variables. These models are their results.

However, the question we have raised is prior to these considerations and concerns the structure preserving character of the mathematisation itself. Simply put, the question is this. Since the output of the OLS models are linear estimates (technically, they are called "expected values")

for the relevant populations, can we be sure the formal statistical assumptions about the structure of the OLS equations map onto reasonable assumptions about the populations of the 66 countries studied and the attitudes their populations might have towards Government intervention? Here we are in precisely the same position Physics is in with regard to its mathematical structures, except Sociology cannot look to a history of successful interventions and implementations on anything like the scale of Physics to justify their use. Instead, what Andersen et al. have to do is to demonstrate professionally recognisable best efforts to shape up their data to reduce any likely mismatch in key areas of their models' assumptions. The possibilities of mismatch fall into three broad clusters: those relating to the form of the relationship among the variables (its linearity); those relating to the statistical independence of the variables; and those relating to the statistical character of the error terms in the regression equations (the 'distance' the values for the attitude of a population actually is from 'expected values' of that population as calculated by the regression function for their country). Since they can know nothing about the error terms, all they can do is make standard assumptions about their form, namely that they are uncorrelated and randomly distributed. And since they know nothing about the actual form of the distribution in the population of the relevant attitudes, they have no grounds on which to prefer a non-linear to a linear form. It follows, the plausibility of Andersen et al.'s analysis has to turn on how they manage what they do know and can affect, their variables. They address this challenge by walking their readers through their analytic procedures.

Analytic Procedures

Operationalisation of the concepts construct a set of relationships, a framework, into which measures on variables can be plugged.

Dependent Variable: Relative support for Government intervention measured by the survey responses.

Independent Variables:

National Level: Level of prosperity measured by GDP per capita; Level of inequality measured by Gini coefficient. These are 'controlled' for cultural factors defined as global region, majority religion, ethnic diversity, level of democracy and communist legacy.

Individual Level: Perception of position in overall income distribution measured by survey response controlled by age, marital status, and religion.

As presented here, the model is a four variable one. For each respondent, the dependent variable is a function of the interaction of independent variables. But each of these independent variables is, in turn, a function of other variables of which there are eight in total. Initially, this gives us 12 degrees of freedom for each respondent. This data space is reduced to 4 by being "controlled". The number respondents included in the analysis is 208,234. The distribution of these respondents in the 4-dimensional Cartesian data space is the data model. We now turn to how this is produced and seen to be produced as managing the potential for an abstraction gap arising out of the interrelationships of the variables. Successfully doing this satisfies the constraint of professional competence.

Strategies of Effacement

When chasing up and down the abstraction ladder, what considerations have to be borne in mind with regard to how 'variables' are managed? There is much professional knowledge and practical experience here and Andersen et al. deploy some salient considerations on the variables they have chosen.

Individual Attitudes

The term "controlled" used with regard to the independent variables is an important signal for a key step that has taken place in the abstraction process. The identified variables are potential "confounders", that is variables which can have interacting effects across a population. Those identified are standard demographic features which are generally recognised as potentially confounding. To remedy this possibility, the data are stratified by the confounding variable (age, marital status etc.) and an analysis of co-variance performed. If there is a statistically significant relationship in the stratified samples of the population, then an adjusting procedure is applied to smooth that out. This is a standard statistical practice and its use displayed simply by reference to the activity of "controlling". Of course, controlling can only be carried out on identified variables. Those listed by Andersen et al, are the ones they take to be material. There can be no guarantee other non-controlled confounding effects remain. One could imagine, for example, educational experience might be relevant and material in the current case whereas height or weight might not. If abstraction is a process of distilling or rendering down, the residues of features not controlled for will remain the in abstracted model. But, once again, this is what every competent mathematical sociologist knows.

In addition to the above possibilities, reliance on 'subjective' measures is known to bring possible complications. Anderson et al. signal their awareness of these by accepting such measures are...

> ....prone to measurement error insofar as respondents may not know their household income or report it accurately. Second, responses to the item depend partly on respondents' perceptions of their countries' household income not just on their actual household income. This circumstance may also induce measurement error because perceptions may be correlated with inequality, with the same absolute income level receiving a higher because perceptions may be correlated with inequality, with the same absolute income level receiving a higher score in a high-inequality country than in a low-inequality country. (Andersen et al, p. 1353)

The problem with such considerations is that although they are well known, there are no satisfactory methods for resolving them. They may be present and material or they may not. There is no way of determining which is which. For the profession, they are well known unknowns.

Prosperity as GDP per capita:

This is defined as the total monetary value of all traded goods and services produced and consumed within a calendar period indexed to a base date. The investigators sum over revenues and costs associated with economic actions. Because of difficulties of comparability, Andersen et al opt to use the UN estimates of national GDP arrived at by pricing a standard basket of goods and services (at PPP or purchasing power parity) indexed to 2005 US dollar values. This index is known to have a number of significant defects. For example, it does not allow for variation in cost structures (e.g., transport and labour costs and sales taxes relevant in different societies) nor the relative cultural value and hence attractiveness or essentialness of some goods and services over others. Both are known to drive price variation (water is regarded as a required public good in most western developed nations and priced as such but not in many developing nations). Not only is GDP not necessarily a measure of experienced prosperity, the use of the PPP method may be prone to an assessable degree of measurement 'slop'.

Income inequality as GINI index:

This index is the difference between two ratios: the normalised cumulative percentage of total income held by a segment of society and the percentage that segment represents of the total population. If 25% of a population receives 75% of the total income, the GINI index is 50%. The absolute limits are 0% and 100%. With 0%, everyone receives the same income. With 100%, only one person receives all the income. As with the GDP measure, there are known pitfalls in using

GINI to make cross national comparisons. For example, there will be clear differences if income is defined simply by earnings rather than including accessible cash as well. Equally, income from the informal sector (the 'Black Economy') is not included. Also excluded from the index are public interventions aimed at reducing inequality such as means tested grants, benefits and subsidies. Finally, there is even the possibility the index will give perverse values. For example, the index may increase indicating increased inequality even though that segment in absolute poverty actually declines.

Political Culture

Clearly, an important factor influencing attitudes to policy implementation is political culture. Andersen et al. include two national level variables to allow investigators to assess its contribution. The first is the existence of democratic institutions. Here the criteria are:

a. Existence of institutions which allow for choice over policy;
b. Institutional constraints on leaders;
c. Constitutional guarantees of civil liberties.

These *pro forma* conditions are taken to index the existence of a democratic culture and coded for any country on a range 0:10. From the discussion, we don't know how the calibration of individual state systems is carried out. Presumably, it is some assessment of the operational effectiveness of the institutions providing the choice, regulating the constraints and guaranteeing the liberties. Since a single value in the range is required for each country, this assessment has to be done by evaluating publicly available and accepted 'objective' distinctions among different liberal Western Democracies and between them and authoritarian regimes of various kinds. This weighing, assessing and judging carries with it similar issues (this time in regard to how the investigators made their evaluations) to those outlined for respondent judgements. As such both are *informal* as opposed to *formal* abstraction processes.

The second variable addresses what might be thought of as the embedding of democratic processes. Here the variable is binary: existence of a communist Government over the last 25 years. This criterion is easy to apply, but what does it represent? Its effect is to act as an index of the history of State Socialism of the Russian, Chinese and related forms. It is not clear, for example, whether the regime in Cuba would count as communist (Cuba is not included in the data) or that of Zambia under Kenneth Kaunda. Equally, no consideration seems to be given to the relative embedding of democratic institutions under non-communist authoritarian regimes such as military juntas and dictatorships.

The indicators used have been developed for large-scale exercises of the kind Andersen et al. are attempting. They summarise their approach to their trade-offs they necessarily have to make as follows.

> If one wants to examine a large number of countries with heterogeneous social characteristics, as we do, these. If one wants to examine a large number of countries with heterogeneous social characteristics, as we do, these issues are unfortunately insurmountable given the current state of cross-national survey research. For example, with the exception of a single case (South Korea in 2018), the WVS lacks an objective measure of income and, with the exception of a single case (South Korea in 2018), the WVS lacks an objective measure of income and, for many countries, lacks other objective measures of economic position, such as social class ..... Faced with a choice between examining a small set of relatively similar countries or a large set of relatively dissimilar countries, we chose the latter because the former has, in our judgment, led to erroneous of relatively dissimilar countries, we chose the latter because the former has, in our judgment, led to erroneous generalizations based on small biased samples. However, the reader should bear in mind the limitations that inhere in our choice. (Andersen et al, p. 1353 reference omitted)

There are, then, well known difficulties with both the 'subjective' and 'objective' dimensions. However, in the absence of a better alternative, to do what they want to do, Andersen et al. have no choice but to use the data they have.

Summary

Andersen and his colleagues provide a thoroughly professional summary of the problem they wish to address and their approach to it. They identify a number of well-known issues with investigations of the kind they have undertaken and outline the decisions they have made in regard to them. In addition, they provide pointers to methods they have used to manage those which they do not discuss. Using these methods, they succeed in reconciling what is initially a somewhat nebulously formulated problem with an amorphous and highly variegated set of data. The problem is given substance and the data is sufficiently simplified and abstract to form a tractable data model.

This success is achieved by a process of property or feature effacement and problem avoidance. However, as we pointed out, techniques of 'confounder control' while professionally applied, are not necessarily exhaustive nor are they necessarily sharply defined. Residues may filter through and relevant properties be inadvertently carved off. Problem avoidance is inevitable in any discipline but especially so in the social sciences where notational ambition quite often

outruns methodological instruments. The lack robust methods (and theories) for tracking the specific processes of attitude and opinion formation at an individual level and then aggregating them to the collective level means the account has perforce to pitch its causal story at the summative level. There are no widely accepted structure preserving 'escalators' available for moving from the models of micro personal attitude formation to the models of macro national or collective attitude formation.[9] But, as we say, this order of property effacement and problem avoidance is well known, thoroughly conventional and normatively regulated. It is the stuff of neophyte training and professional off-the-record discussions. Its use constitutes much of the taken for granted background read into research reports by professionally competent readers.

Section 3. Conclusion

Our aim has not been to justify the use of mathematics in the natural and social sciences by providing an answer to the philosophical puzzle with which we started. Neither do we wish to argue for its rejection. What we have tried to do is show how competent disciplinary good practice manages the problem of relating mathematical structures, data structures and phenomenological experience and how it displays that management in its research reporting. We called this the double-fitting of data and phenomena. The degrees of freedom in any such fitting are regulated by the norms of a discipline's own practise. In the natural sciences, these norms impose pretty austere requirements on a description and its notation. In the social sciences, for reasons which we have touched on here and elsewhere, the requirements are more loose-limbed. It is for each discipline to choose the constraints under which it wishes to work and such choices will be made with the need to ensure investigations are "doable" in Leiberson's [1987] sense of not imposing impossible-to-satisfy objectives simply out of a desire to be 'more rigorous' or 'more scientific'. Double fitting phenomena and data is not a defect in either the natural or the social sciences. It is an ineradicable fact of professional life, a "normal natural trouble", to use Garfinkel's phrase, to be handled as best one can in one's investigative reporting. The management of the abstraction gap is not a surreptitious, sub-rosa blemish on investigative disciplines but a routine feature of their own lived experience.

All of which takes us back to Nguen and Frigg's questions. With regards to the second, the justification of a particular mapping, the compromises, carried through residues, workarounds

[9] Behr and Passerini [Bahr and Passerini 1998a], [Bahr and Passerini 1998b] have suggested Mean Field Approximation methods adapted from Thermodynamics might be suitable. Thus far, their proposals have not really been taken up.

and avoidances Andersen et al. were required to accept to get their investigation to work, seem to suggest to existence of not an insubstantial abstraction gap between their data model and the social phenomenology which is its target. How substantial the gap might be is impossible to assess with current methods. All we can do is turn back to the argument Euler, the paramount mapper of Classical Mechanics, used to justify the extension of the mathematical to the physical, namely the usefulness of the results. But, in our experience, to ask how useful Andersen et al.'s study is or might be, or indeed how useful any application of mathematics to the social might be, is not liable to resolve anything. More likely it will just add confusion and dispute.

As for the first question, given the applications of Mathematics in Sociology almost always use functions on $\mathbb{R}^N$ or some related number system, the security of such mapping rests on the security of the foundations of $\mathbb{R}^N$ or whatever as a system of mathematical objects. The trouble here is there is just as much dispute on this as there is in Sociology over the applicability of its research findings. Today, the foundations of the number systems are sought by looking for a new setting, a geometrical conception of number spaces in which they can all be placed. But achieving that conception requires the invention of new methods and with them new objects, such as imaginary numbers, vanishing points, lines at infinity and so on which make the mathematics work but whose 'reality' many mathematicians balk at. As Mark Wilson [Wilson 2001] put it, the search for new and better theorems and proofs is a process of creation and innovation which is in permanent tension with observance of the primary mathematical virtue of certainty.

As for the third question, how mathematical structures can *explain* the non-mathematical and especially the social, it is probably best to let that sleeping dog lie.

Bibliography

Andersen, R., Curtis, J., and Brym, R. 2021. Public Support for Social Security in 66 countries. British Journal of Sociology 72, 5, 1347–1377.

Bahr, D. and Passerini, E. 1998a. Statistical mechanics of opinion formation and collective behaviour: Micro-sociology. The Journal of Mathematical Sociology 23, 1, 1–27.

Bahr, D. and Passerini, E. 1998b. Statistical Mechanics of collective behavior:: Macro-sociology. The Journal of Mathematical Sociology 23, 1, 29–49.

Baldamus, W.B. 1971. Typres of Trivialisation. Faculty of Commerce, University of Birmingham, Birmingham.

Frigg, R. and Nguen, J. 2019. Mirrors without Warnings. Synthese 198, 6, 2427–2447.

Leiberson, S. 1987. Making it Count! The Improvement of Social Research and Theory. University of California Press, Berkeley, Ca.

Livingston, E. 1986. The Ethnomethodological Foundations of Mathematics. Routledge & Kegan Paul, London.

Livingston, E. 1999. Cultures of Proving. Social Studies of Science 29, 6, 867–888.

Nguen, J. and Frigg, R. 2022. Scientific Representation. Cambridge University Press.

Pickering, A. 1995. The Mangle of Practice. University of Chicago Press, Chicago.

Shapiro, S. 2000. Thinking about Mathematics. Oxford University Press, New York.

Stemeroff, N. 2021. Structuralism and the conformity of mathematics and nature. Studies in History and Philosophy of Science, Part A [C], 86, 84–92.

Wilson, J. 2010. Non-reductive Physicalism and Degrees of Freedon. The British Journal of Philosophy of Science 61, 2, 279–311.

Wilson, M. 2001. The Imitation of Rigor. Oxford University Press, Oxford.

Woodward, J. and Bogen, J. 1988. Saving the Phenomena. The Philosophical Review XCVII, 303–352.