![JASSS logo]

Corinna Elsenbroich (2012)

Explanation in Agent-Based Modelling: Functions, Causality or Mechanisms?

Abstract

What kind of knowledge can we obtain from agent-based models? The claim that they help us to study the social world needs unpacking. I will defend agent-based modelling against a recent criticism that undermines its potential as a method to investigate underlying mechanisms and provide explanations of social phenomena. I show that the criticism is unwarranted and the problem can be resolved with an account of explanation that is associated with the social sciences anyway, the mechanism account of explanation developed in Machamer et al. (2000). I finish off discussing the mechanism account with relation to prediction in agent-based modelling.

Keywords:
Philosophy of Social Science, Causal Explanation, Functional Explanation, Mechanism Explanation, Analytic Sociology

Introduction

1.1 Agent-based modelling (ABM hereafter) has, over the past three decades, become a more and more influential method in the social sciences (Gilbert 2008) with applications in sociology (Macy and Sato 2002, Gilbert et al. 2010, Hamill and Gilbert 2009), criminology (van Baal 2004, 2008) and even some in economics (see Tesfatsion 2003 for an introduction, Gilbert et al. 2008 for an application and Buchanan 2009 for a justification).

1.2 Given this growing interest in ABM it is important to assess the kind of knowledge that can be obtained from ABM. Although there exists a body of literature on the epistemology of simulation (e.g. Humphreys 2002, Humphreys and Bedau 2007, Humphreys 2009, Winsberg 2003, Frigg and Reiss 2009) much of it is concerned with simulation in the natural sciences, possibly stretching to economics but rarely to the less mathematical of the social sciences.

1.3 Recently ABM has come under philosophical attack. Grüne-Yanoff (2009) argues that ABM, against the claims of modellers, is not able to provide causal explanations of social phenomena. Using a prominent example of an artificial society (Dean et al. 1999), he intends to show up a host of problems that beset ABM. His main argument is that ABM can only ever provide *partial explanations*. But a causal explanation has to be complete, i.e. tell the 'whole' causal (hi)story of an event, and no ABM is able to provide this kind of full causal history. He uses an account of a functional explanation from Cummins (1975) to provide a positive account of what kind of explanation can be provided by an ABM.

1.4 In this article I will show that Grüne-Yanoff's argument is flawed in several ways. The problem of partial explanations and all the ensuing problems discussed below as problems of ABM hold for at least all of the social sciences, possibly all sciences. Grüne-Yanoff proposes an alternative kind of explanation, a *functional explanation*, which is epistemically second class. A relatively recent account of explanation, the mechanism account (Machamer et al. 2000), elegantly resolves the problem of causal explanations having to be complete. After showing that Grüne-Yanoff's position is not specific to ABM and that it is unnecessary to resort to a second class explanatory account, I use the mechanism account of explanation to resolve another recent debate on explanation and prediction in ABM (Epstein 2008, Thompson and Derr 2009, Troitzsch 2009).

The Goal of the Social Sciences

2.1 Before heading into the question of what kind of explanation we can expect from ABM I want to briefly discuss different kinds of

explanation in general and kinds of explanation in the social sciences.

What kind of Explanation?

2.2   The starting point of a proper philosophy of explanation is often put in the 1940 with Hempel and Oppenheim's (1948) account of the covering law model of explanation. It is the creation of a logic of explanation, defining an explanation as a deduction of the explanandum from a set of general laws and initial conditions. Importantly, the covering law model sees explanation and prediction as symmetric concepts. For indeterministic events the inductive-statistical model of explanation is invoked where the explanandum is rendered highly likely as a conclusion of the general laws plus initial conditions. There are many objections to this model of explanation, e.g. the purely formal nature of the account, asking questions of relevance and meaning, i.e. one might be able to deduce an explanandum from a set of premises without it constituting an explanation. Another widespread attack is the symmetry thesis of the account that sees explanation and prediction as having the same formal structure, the only difference being their temporal orientation.[1]

2.3   Since the 1940s there has been a proliferation of models of explanation, e.g. the pragmatic account developed in van Fraassen (1980) or accounts of explanation as serving understanding rather than explanation (de Regt 2009). In the following discussion I limit myself to accounts that allow for a realist ontology as we only need to consider accounts stronger in terms of realism than the functional explanation account proposed in Cummins (1975). Realist accounts of explanation can be roughly classified as the following (Douglas 2009).

1. *Covering Law Explanations*: the deduction of the explanans from general laws and initial conditions (Hempel and Oppenheim 1948). Example: explaining the trajectory of a projectile using Newtonian physics.
2. *Causal Explanations*: relaying a step-by-step account of how an event came about, identifying all relevant facts pertaining to the event (Salmon 1998). Example: explaining the occurrence of a car accident.
3. *Mechanistic Explanations*: a kind of causal explanation where no full causal story can be told. Entities and their causal connections are identified to give an explanatory account of a phenomenon (Machamer et al. 2000). Example: explaining the recent financial crisis by sub-prime mortgage lending.
4. *Unifying Explanations*: to show that the explanans falls under a general theory that explains many similar phenomena (Kitcher 1976). Example: evolutionary explanation of explananda such as the peacock's tail and the human brain.

2.4   The general law in the covering law account of explanation has the function to make the link between one state S_1 (the initial conditions) and another state (S_2) (the explanandum). The law 'covers' the transition between S_1 and S_2. If we take the covering law approach as the starting point of a theory of scientific explanation the other three accounts can be related to it as relaxing the general law requirement in different ways.

2.5   In the case of causal and mechanistic explanations, the relaxation happens by no longer using a general law to justify the transition from S_1 to S_2 but by proposing a direct causal connection between S_1 and S_2. For the traditional causal explanation a full causal story of an event has to be given, i.e. S_1 must lead to S_n via $S_2 \ldots S_{n-1}$. In a mechanistic explanation the causal powers are proposed to be directly connected to the entities involved in the event (e.g. via 'capacities' or 'activities'). No full causal story needs to be told but the entities must be shown to be 'at work'. Unifying explanations relax the lawlikeness of the theoretical cover by replacing it by a general theory linking S_1 and S_2.

2.6   As stated above, decades of criticism have rendered the covering law model of explanation a nice idea without much application. Most sciences do not have the laws of nature necessary to obtain an explanation and prediction symmetry. The main problem with the causal model of explanation is how to actually distinguish a causal process from any other correlation process in particular in complex systems. The problems of explanatory relevance and explanatory asymmetries[2] are a result of an impoverished ontology of an account of causation which shies away from attributing causal powers to entities and tries to cover causal connections by regularities. I discuss mechanism explanations in the next section as the most appropriate kind of explanation for the social sciences. The unification theory of explanation is a useful theory in some contexts but has the problem of generality of an explanation leading to irrelevance (Lipton 2004). Let us now look at explanation in the social sciences.

What kind of Explanation for the Social Sciences?

2.7   There are different conceptions of the ontology and epistemology of the social world. One such conception is that of analytic sociology in which the social sciences are an investigation into the mechanisms underlying social phenomena (Elster 1989, Hedström and Swedberg 1998). The mechanism approach is in opposition both to hermeneutic approaches, which claim that the social world cannot be explained but only interpreted and to statistical explanations based on the covering law approach discussed above.[3] Seeing the social sciences as concerned with mechanisms means to not allow "black-box explanations" such as statistical correlations. Although statistical correlations can be used as evidence for causal associations, they are not an explanation in themselves as they do not lay open the "cogs and wheels" operating to produce the phenomenon in question.

2.8   There are several different definitions of what mechanisms are (for a recent review see Hedström and Ylikoski 2010). A mechanism approach neither reduces social entities to physical entities nor sees social mechanisms as the same as physical mechanisms. Hedström and Ylikoski state that an adequate account of mechanisms in the social sciences needs a proper theory

of action. The authors argue that a mechanism-based social science, although often associated with rational choice theory, is in fact a step away from the empirically false assumptions of rational choice. In order to uncover mechanisms it is not enough to "save the phenomena" (Duhem 1954) but assumptions need to have empirical foundation. On the other hand, mechanism based approaches can be very abstract and start from very simple assumptions and specifications.

2.9   The assumptions of a mechanism view of the social world can be summarised as:

     i.   causality is more than mere regularity,
    ii.   there are entities at work producing social phenomena and
   iii.   we can identify these entities and their interactions and thus find the mechanisms underlying social processes.

An account of mechanism explanations can be found in Machamer et al. (2000). A further account on causal explanations in the social sciences is developed in Yilkoski (2001).

🌎 Functional Capacities or Causal Explanations?

3.1   In this section I look at a recent publication highly critical of ABM. Grüne-Yanoff (2009) argues that ABM has been misconstrued by modellers as explaining causal histories of social phenomena. He uses the Anazasi model from Dean et al. (1999) and Axtell et al. (2002) for his discussion. The model describes the population dynamics in the Long House Valley, Arizona. The Anazasi are a population that inhabited the Long House Valley, starting slowly from the introduction of a maize production in BC 1800 until AD1300 at which point there was a sudden and complete population exodus from the area. The landscape of the area is reconstructed in the model, using paleo-geographical data. The agents in the model are households. Agents are defined by attributes like Maize consumption, Maize production, storage etc. There are two reasons for households to move: marriage (household fission) and insufficient harvest (household relocation). Using a host of other paleo-environmental data, the population dynamics are modeled for the timespan between AD 800-1300, a period, which is relatively data-rich, relatively in terms of archeology. Runs of the model give outputs of dynamics that are very close to the historic dynamics curve. The total exodus is not replicated in any run, however. The authors conclude that

> To "explain" an observed spatio-temporal history is to specify agents that generate or grow this history. By this criterion, our strictly environmental account of the evolution of this society during this period goes a long way towards explaining this history. (Axtell et al. 2002, p.7279)

3.2   Grüne-Yanoff argues that this simulation does not give a causal explanation of the population dynamics of the Anazasi. The argument is not that the simulation is not an explanation of the Anazasi population but that it is not a *causal* explanation. The simulation does not tell a full causal history of the population dynamics, for example, the total exodus is not an outcome of any run of the simulation, thus leaving the explanation partial. Grüne-Yanoff argues that there is no such thing as a partial causal explanation. A causal explanation has to give a full account of all interactions leading to a phenomenon as there is no formal criterion to distinguish bad partial causal explanations from good partial causal explanations. As we have seen above, a causal explanation indeed needs to provide a full causal history as it cannot rely on any overarching theory to cover connections between events. We discuss Grüne-Yanoff's criticisms one-by-one in the following subsections showing that they are not problems unique to ABM.

3.3   Grüne-Yanoff goes on to discuss what kind of explanation ABM can provide, concluding that the most appropriate theory of explanation for ABM is that of a *functional explanation* from Cummins (1975). In the following sections I discuss Grüne-Yanoff's arguments why ABM can only provide partial explanations to then discuss in more detail the functional explanation account and how it supposedly solves the explanatory problem of ABM.

3.4   Grüne-Yanoff's position has been challenged elsewhere on other accounts (Chattoe-Brown and Elsenbroich 2011) so here I only discuss those points relevant for the discussion of the explanatory power of ABM.

Data-problems

3.5   Grüne-Yanoff identifies two problems concerning data in the Anazasi Simulation contributing to it being merely a partial explanation. The first problem has to do with the input data. In a simulation of a car in a car crash we obtain a result exemplifying a causal explanation (e.g. why the windscreen smashed). In order to simulate the car and the crash we have full information of the material and the structure of the car. We also know all the causal laws at work in a crash situation and know how the constituent parts interact. In the case of the Anazasi simulation we do not know all the micro-mechanisms at work, i.e. the behavioural rules of the agents etc. nor do we have any overarching laws. Thus, we will not have a full explanation of any macro-phenomenon. The second problem resides with the output data. The simulation of the Anazasi replicates the population dynamics curve quite well. Both the closest run, presented in (Axtell et al. 2002, p. 7278), and the average run taken from about 15 replications show a very close fit. However, no run has ever replicated the complete exodus of the population in 1300. Thus, the explanation cannot be a complete explanation, as at least one real life phenomenon is not explained by it, Grüne-Yanoff argues. There are two problems with this analysis:

    1.   The car crash simulation is not intended to give an explanation but intended to test whether the car is safe. The Anazasi

simulation is intended to replicate a set of output data by making a set of assumptions.

Grüne-Yanoff's account makes any abductive inference from simulation impossible. It also makes abductive inference for a causal explanation impossible. At any point where a competing explanation exists, Grüne-Yanoff's account necessarily reduces the explanation to a functional explanation and ergo it makes causal explanation impossible in the social sciences. This is not a problem of ABM but of social science as a whole.

2. That the input data is unreliable is a feature of the social sciences, not of the methodology of simulation. It is almost common sense that a model based on archeological data will be missing some factors. In addition, in the social sciences in general we cannot isolate experimentally micro-level factors. Any data in the social sciences will be partial, possibly unreliable, and full of problems endemic to the social sciences as a whole. Insisting on complete knowledge of micro-phenomena for a causal explanation makes causal explanation impossible in the social sciences. This is also not a problem of ABM but of social science as a whole.

What Grüne-Yanoff identifies as data problems of ABM are simply data problems of the social sciences. It is not impossible to defend a position in which the social sciences do not aim to find causal laws (e.g. interpretative and descriptive approaches, see for example Rabinow and Sullivan 1979) but from such a position it makes no sense to attack ABM to not come up with the goods as there is no goods to come up with.

Falsifying and Curve Fitting

3.6   The Anazasi model is compared to a climate model, discussed in Küppers and Lenhard (2005). Initially the climate model was run based on simple equations, known to be true about the system; it 'exploded' after a short run. The simulators then deliberately added some false assumptions, e.g. that the kinetic energy in the atmosphere remains constant. By adding this assumption the system ran smoothly and gave appropriate retrodiction of climate data, similar to what the Anazasi simulation does. Grüne-Yanoff states that a simulation with false assumptions cannot enhance our causal understanding of the world but it is still useful as it helps us to understand processes functionally.

3.7   Again, the climate simulation and the Anazasi simulation are inherently different. Both simulations retrodict data but for different purposes. Whilst Axtell et. al. are interested in seeing in how far data available on environmental factors can explain the population dynamics in the Long Valley, the climate simulation uses the retrodiction as a proxy for the capacity of the simulation to predict. The falsification in the climate simulation is used to make the simulation run in the first place and to fit the data. No conscious falsification has been performed in the Anazasi model to fit the historic data curve. The assumptions in the model might be incomplete but they have not been *falsified* to fit the model to the data.

Level-Distinctions

3.8   The final argument against ABMs resulting in causal explanations is about levels of explanation. Towards the end of the paper Grüne-Yanoff makes a distinction between functional and causal explanations involving different levels of phenomena. Functional explanations use constituent capacities (lower level) to explain system capacities (higher level). Causal explanations, in comparison, need to explain phenomena on the same level. For the Anazasi simulation Grüne-Yanoff states that a dispersion (explanandum) is supposed to be explained by the individual movings (explanans) but then he identifies the disperson and the movings, meaning they cannot be on a different level.

The dispersion is nothing but the individual movings.(Grüne-Yanoff 2009, p. 552)

Although it can be argued that a dispersion is not just the individual movings as it is a coordinated movement, the following example exemplifies a clearer case of micro and macro levels. Let us have a look at a well-known ABM, the Schelling Model of Segregation from Schelling (1971). Schelling wondered about the phenomenon of segregation in American cities. To explore the phenomenon he devised a simple model. Imagine a grid with agents randomly allocated to the patches on the grid. The agents come in two colours, red and green. Agents have a threshold for how many agents of the other colour they tolerate in their neighbourhood. If the number of other coloured agents exceeds that threshold, the agent moves to an arbitrary other patch on the grid. Schelling initially executed this set of agent interactions on a checkered board by hand. He found that clustering of colours and hence segregation resulted even with agents having rather weak preferences for similarity.

3.9   Clustering/segregation comes about due to the movements of agents according to their preferences. Here the thresholds cause agents to move (same (micro) level) the movings according to those preferences cause clustering/segregation (higher (macro) level). It is only a partial explanation of the real world phenomenon as no real world segregation is simply caused by the preference relation; many other variables play into real segregation. But would we say that the segregation is nothing but the individual movings? There seems to be a more interesting process explored in this simulation experiment.

Limit-Problems

3.10   Grüne-Yanoff states that a full functional explanation is a causal explanation. This uniqueness is explicitly used in Cumins' Individuation criterion (below). The problem is that we will never have a causal explanation as we can never really be sure we have identified all the real entities and mechanisms at work. The theory of functional explanation does not allow a part causal part

functional explanation with the possibility of an increasing causal part in the face of additional information. It also does not allow for isolating causes for phenomena. This would mean that the Shelling model of segregation is not a causal explanation of segregation at all. It is clear that the Shelling model does not tell the whole story of segregation but it shows that segregation *can be caused* by the possibility of movement even at very high tolerance thresholds.

Functional explanation

3.11 So Grüne-Yanoff is right in his assessment that ABM does not provide full causal histories and thus cannot provide causal explanations (if the only view of causal explanation is an account relying on the regularity view of causality). He goes on to discuss an account of explanation allowing ABM to save (explanatory) face. I discuss the account below and argue that the account, rather than providing a second class explanation, as is suggested in Grüne-Yanoff (2009), coincides almost exactly with an explanatory account devised for complex phenomena instead of a strict causal account, the mechanism account (Machamer et al. 2000).

3.12 According to Cummins (1975), the explanandum in a functional explanation is a system S ψ-ing, i.e. displaying the capacity ψ. The explanans of a functional explanation has three parts:

   i. an analytic account A of a system S ψ-ing which contains
   ii. the description of a component x φ-ing and
   iii. the claim that x indeed can φ.

3.13 The advantage of this functional explanation using the notion of *capacity* rather than *cause* as a potential or possible explanation is threefold:

   1. *Individuation* according to possible functions rather than factors or mechanisms. Why this is important seems to be a difference in specificity, i.e. a factor or mechanism is a unique instantiation whereas a function can have multiple instantiations. So, we find out something like 'there is *something* that performs *some* dampening' without knowing exactly what entity or collection of entities produces the effect.
   2. *Transferability* across different causal structures. Grüne-Yanoff uses the example of the Ising Model of statistical mechanics. The model can be applied to ferromagnets as well as to market dynamics. The only way the same model can be applied to two different systems which clearly do not share a set of entities is that they have something else in common, something like structure, here called *functional organisation*. Functional organisation 'can be analysed with the same model and this model may improve our understanding of how each system acquires the capacities it has through the interactions of its subsystem.' (Grüne-Yanoff 2009)
   3. *Level-distinction* in functional analysis means that the analysis shows how lower level capacities constitute higher-level capacities. The differentiation is between constitutive relationships at different levels in a functional explanation and causal relationships at the same level. (See §3.10 above.)

3.14 First of all I think the terminology of the above rather unfortunate. A functional explanation, as commonly used in the social sciences, is ontological and a *sui generis* feature of the social world. There are plenty of problems besetting functional explanations, among them their teleological nature and their ascription of intentions to non-intentional entities. Although feathered wings help birds to fly, wings did not (over time) become feathered to enable birds to fly. In the social sciences there are many functional explanations of this kind in which phenomena are explained by their 'benefit to society' (e.g. Parsons' Structural Functionalism). For an account of the *sui generis* kind of functional explanation and ABM see Chattoe-Brown (2006).

3.15 The notion of functional explanation above is epistemological in the sense that we can find out the functions of different parts within a system. It is a partial causal explanation viable only as an interim solution whilst we do not know better given that partial causal explanations are impossible as there are no quality criteria for partial causal explanation. The truth of a causal explanation can only be assessed by knowing each causal step. Thus a functional explanation *à la* Cummins is a second-class explanation.

3.16 Secondly, and more importantly, this capacity account of explanation can be mapped onto the mechanism account of explanation advocated in Machamer et al. (2000) as a causal explanation for complex systems. Machamer et. al. propose that the ontology of causality as regularities makes capturing causal processes in complex systems impossible. Instead they propose an ontology that involves *entities* and *activities*. In the above account, x is an entity in the system S that displays the activity of ψ-ing leading to the system S ψ-ing. Committing to an ontology of causal capacities of entities or causal activities of entities, Cummin's account of functional explanation becomes a mechanism explanation *a la* Machamer et. al. This means it is not a second-class explanation at all but a causal explanation in an ontology that sees causality as more than mere regularity and the kind of causal explanation advocated for complex systems of which society can be seen as an example. Thus again, the partial nature of explanations in ABM is not unique to ABM but a feature of the social sciences.

3.17 Furthermore, the problem of partial causal explanations in the social sciences is well known, resulting in varied proposals of accounts of explanations. Some of these accounts eliminate a realist ontology altogether, like van Fraassen's (1980) pragmatist account or de Regt's (2009) account of explanation as understanding. I opted for an account that preserves a realist ontology but is also adequate for systems such as social systems. This account is the mechanism account of explanation (Machamer et al. 2000) because it a) renders explanations from ABM causal rather than the second class functional explanations proposed by Grüne-Yanoff, b) it fits in with a coherent interpretation of the social sciences as uncovering mechanisms underlying social

phenomena c) the mechanism account of social science is closest as an aim to the methodology of ABM and finally, d) the account of mechanism explanations helps to resolve another epistemological/methodological debate in ABM, the problem of explanatory power with limited predictive power.

🌍 Predictions and Explanations

4.1 In this section I discuss the problem of prediction in ABM. In the first principled study of explanations in the philosophy of science, explanation and prediction were inextricably linked (Hempel and Oppenheim 1948). This link is a logical equivalence between prediction and explanation called the *symmetry thesis* stating that the logical structure of both explanation and prediction is the same and the only difference between them is the temporal orientation (prediction is future oriented whereas explanation is about the past). Essentially, if we can predict a phenomenon to occur, we know what brings it about and thus have explained it. As discussed above, this equivalence resulted from a specific model of explanations, the covering law model. In the covering law model, an explanation is the deduction of an explanans from a (set of) general law(s) plus (a set of) initial conditions.

4.2 Prediction is a thorn in the side of ABM. Models claim to replicate processes of the real world but no detailed prediction has ever been derived from an ABM.[4] One important reason why ABM are expected to predict is that traditionally, prediction is what simulations are for. Two examples of simulations were discussed in earlier sections, the crash-test car simulation and the climate model. The car crash simulation is devised solely to simulate the exact procedure of the car crash, thus predicting the exact impact of the crash on the car. Its purpose is to predict so that the real world car can be improved. This predictive power is achieved by the full knowledge of the causal interconnections of the car (of its structure, its materials, etc.).

4.3 The climate model also started off as a truthful mechanistic system. The initial programming was according to the laws supposedly underlying the system. The simulation 'exploded', rendering the truthful representation useless for prediction. Some falsifying assumptions were made until finally the simulation retrodicted historic data adequately. The reason for the simulation is to predict future climate development. Given the falsified assumptions we do not believe this prediction due to the exact replication of the processes in the climate system as was the case in the crash test dummy example. Here we believe the prediction because of the fit of historic data, its successful retrodiction. We could call this an inductive prediction. The difference between inductive and causal prediction is well made in Troitzsch (2009).

4.4 The identification of prediction and explanation has long been debated (see e.g. Rescher 1957) resulting in a proliferation of explanatory models divorced from prediction. Explanation learned to stand on its own feet as a sole legitimate goal of science with prediction relegated to the fringes.

4.5 Causal, mechanistic and unifying explanations have been divorced from prediction by their singularity (causality), their complexity (mechanism) and their generality (unification). But can an explanation really be explanatory without allowing prediction at all?

4.6 In a recent discussion of the epistemology of ABM, Epstein (2008) deliberately cuts the cord between prediction and explanation, saying that there are many other reasons to model. He is certainly right in saying that prediction is not the sole justification of a model, in particular prediction as commonly defined in the natural sciences, i.e. predicting at a pretty detailed level what happens to a system or parts of a system in the future. Against this rather defensive position, Thompson and Derr (2009) argue that explanation and prediction must not and need not be divorced, only the notion of prediction needs to be widened to include more general predictions, such as "earthquakes happen" as a prediction of tectonic plate theory. Troitzsch (2009) makes an important distinction between three levels of specificity of predictions. He also distinguishes between stochastic and deterministic predictions. His solution to the debate on prediction and explanation is that any good explanation will yield at least a prediction of type one:

> Which kinds of behaviour can be expected [from a system like this] under arbitrarily given parameter combinations and initial conditions? (Troitzsch 2009, 1.1)

Sometimes it will even yield an explanation of type two:

> Which kind of behaviour will a given target system (whose parameters and previous states may or may not have been precisely measured) display in the near future? (Troitzsch 2009, 1.1)

It is, however, not necessary for an explanation to provide a prediction of type three:

> Which state will the target system reach in the near future, again given parameters and previous states which may or may not have been precisely measured? (Troitzsch 2009, 1.1)

Troitzsch's reply to Thompson and Derr's weakening of the notion of prediction to the prediction "earthquakes occur" from plate tectonics is that humanity will know that fact after experiencing the first earthquake without needing any theory at all.

4.7 For Troitzsch the problem of prediction and explanation is a purely epistemic one of not knowing the initial conditions of a system well enough to warrant prediction. Troitzsch concludes that the symmetry thesis of explanation and prediction is alive and well "from a logical point of view but not from a practical point of view" (Troitzsch 2009, 1.6). But is this really all there is to it?

4.8 The interpretation of the symmetry thesis dividing the logical and epistemic level states that explanations and predictions are

logically equivalent but due to a lack of knowledge of initial conditions we might have non-predictive explanations. It does not answer the problem of non-explanatory predictions.

4.9 Let us take the climate simulation from Küppers and Lenhard (2005) again. The processes in the simulation were deliberately falsified. The simulation retrodicts climate data well and we infer from this capacity to retrodict the capacity to provide prediction of type 2 above. But is this kind of prediction explanatory?

4.10 Let us imagine a scenario: There is a competition to draw an imaginary climate curve up to the timespan between 2000 BC to 3000 AD. Let us presume we match the curves to the real data and then select the one that is the closest fit in terms of retrodiction. Would we be in any way justified to presume that the future prediction is true? We clearly would not. Does it mean that it is necessarily false? Certainly not, but a match would be pure luck (or some supernatural clairvoyance on the part of the curve maker).

4.11 We might have prediction without the underlying mechanism but no reason to believe it is true. We certainly do not have an explanation of climate development.

4.12 The problem of non-explanatory predictions has been discussed widely in the philosophical literature. In Salmon's (1978) "Why ask, 'Why?'?", Salmon argues that if we had all the information in a deterministic world and would be able to exactly predict any future states of the world, we would still want more, we would still want an answer to a Why-question. The reason why we would still want this answer is that we would want knowledge of the underlying mechanisms bringing about the future states.

4.13 Although it is questionable whether we would really still want answers to Why-questions if we were able to predict everything perfectly, in a world where we *cannot* predict everything perfectly, answers to Why-questions providing knowledge of the underlying mechanisms are essential because it helps us to predict some bits of the world a little bit better. Douglas (2009) argues along those lines that the symmetry thesis does not give an adequate account of explanation but neither does any account leaving out prediction altogether. For her explanations are tools to help us generate predictions about the world by providing conceptualisations of the world.

> Explanations help us to organize the complex world we encounter, making it cognitively manageable (which may be why they also give us a sense of understanding). (Douglas 2009, p. 454)

This cognitive account of explanation and prediction together with Troitzsch's different layers of system state prediction can give us a handle on ABM.

4.14 Let us look at an example. In Gilbert et al. (2008) an agent-based model of the English housing market is developed. In this model there are agents acting as buyers, sellers and realtors (estate agents) according to different goals and intentions, leading to different roles in market interactions. The environment is a grid with houses of initially random values. The agents also have different incomes and savings and are on different steps of the housing ladder (e.g. first time buyers). The model implements a host of economic variables, such as interest rates, inflation, Gini index etc. One of the most interesting outcomes of the model is that the level of interest rates has much less influence on house prices than for example the number of first time buyers entering the market. The model was published in 2008 and shortly afterwards the banking crisis was followed by a (slight) housing crash; and the lack of first time buyers contributed to this crash as banks suddenly wanted a large deposit for mortgages, at least for mortgages at reasonable interest rates, resulting in many people no longer being able to afford to enter the market as first time buyers leading to many properties staying on the market for longer until sellers reduced the price. The model of the housing market did not predict a housing market crash due to a lack of first time buyers but it helps to investigate the influence of different parameters, such as interest rates, first time buyers, etc. on house prices.

4.15 Mechanism explanations give us an account of how a system works, a set of patterns in particular contexts. By invoking the entities and activities at work, we can work out what will happen in similar contexts containing such entities and activities. They help us to generate predictions which can be tested in experimental settings and, if the predictions are false, the mechanism framework requires one to figure out which mechanism was at fault (see Douglas 2009 for a discussion on mechanism explanations and prediction).

🌍 Conclusion

5.1 In this article I have argued against a criticism recently launched at ABM of not being able to provide causal explanations in Grüne-Yanoff (2009). My main arguments against his position are that his criticisms are not specific to ABM but beset the whole of the social sciences and that the account of explanation, he sees as appropriate for ABM, is not the second rate explanation he makes it out to be but indeed the appropriate account when used with a sensible ontology for the social sciences. As a result ABM are a methodology providing mechanism explanations, a kind of causal explanations for the social sciences (Machamer et al. 2000). I have further argued for mechanism explanations resolving the problem of prediction and explanation (Epstein 2008, Thompson and Derr 2009, Troitzsch 2009).

5.2 Not all problems are solved by ABM providing mechanism explanations. The ontology of entities and activities proposed in Machamer et al. (2000) is by no means uncontroversial. A smaller ontology of entities and capacities (advocated for example by Cartwright 1989) is often preferred to the double commitment above. Although the ontological question is important for the

philosophy of science, the account of mechanism explanations can stand without its resolution. Whether activities have ontic status or whether they are outcomes of capacities, they can feature in an explanation of a complex phenomenon.

Notes

[1]For further discussion of criticisms, see Salmon (1990).

[2]For explanatory relevance consider the explanation of birth control pills 'preventing' pregnancies in men, for explanatory asymmetry imagine a flagpole of height h with a shadow of length l, h can explain l but it makes no sense to explain the height of the flagpole by its shadow (both examples are from Salmon 1990).

[3]See Hempel (1965) for the inductive-statistical explanation.

[4]In April 2009 Scott Moss asked the SimSoc email distribution list whether there as ever been an ABM making a successful prediction of a policy implementation. The answer was negative, leading to a discussion on the same list headed 'What's the point?'

References

AXTELL, R., Epstein, J. M., Dean, J. S., Gumerman, G. J., Swedlund, A. C., Harburger, J., Chakravarty, S., Hammond, R., Parker, J., and Parker, M. (2002). Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. *PNAS*, 99: 7275-7279. [doi:10.1073/pnas.092080799]

BUCHANAN, M. B. (2009). Meltdown modelling. Could agent-based computer models prevent another financial crisis? *Nature*, 460(7256):680-682. [doi:10.1038/460680a]

CARTWRIGHT, N. (1989). *Nature's Capacities and their Measurement*. Oxford.

CHATTOE-BROWN, E. (2006). Using simulation to develop testable functionalist explanations: a case study of church survival. *British Journal of Sociology*, 57(3):379-397. [doi:10.1111/j.1468-4446.2006.00116.x]

CHATTOE-BROWN, E. and Elsenbroich, C. (2011). The explanatory potential of philosophical analysis: A comment on Till Grüne-Yanoff. http://www2.le.ac.uk/departments/sociology/documents/tgyexplan.doc.

CUMMINS, R. (1975). Functional analysis. *Journal of Philosophy*, 72(20):741-765. [doi:10.2307/2024640]

DE REGT, H. W. (2009). The epistemic value of understanding. *Philosophy of Science*, 76(5). [doi:10.1086/605795]

DEAN, J. S., Gumerman, G. J., Epstein, J. M., Axtell, R., Swedlund, A. C., Parker, M. T., and McCarroll, S. (1999). Understanding Anasazi culture change through agent-based modeling. Working papers, Santa Fe Institute.

DOUGLAS, H. E. (2009). Reintroducing prediction to explanation. *Philosophy of Science*, 76(October):444-463. [doi:10.1086/648111]

DUHEM, P. (1954). *Aim and Structure of Physical Theory*. Princeton University Press. [doi:10.1119/1.1933818]

ELSTER, J. (1989). *Nuts and Bolts for the Social Sciences*. Cambridge University Press. [doi:10.1017/cbo9780511812255]

EPSTEIN, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, 11(4):12 http://jasss.soc.surrey.ac.uk/11/4/12.html.

EPSTEIN, J. M. and Axtell, R. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings Institution Press.

FRIGG, R. and Reiss, J. (2009). The philosophy of simulation: Hot new issues or same old stew. *Synthese*. 169(3):593–613. [doi:10.1007/s11229-008-9438-z]

GILBERT, N. (2008). *Agent-Based Models*. Quantitative Applications in the Social Sciences, 153. Sage Publications.

GILBERT, N., Ahrweiler, P., and Pyka, A. (2010). Learning in innovation networks: Some simulation experiments. In Ahrweiler, P., editor, *Innovation in Complex Social Systems*. London: Routledge.

GILBERT, N., Hawksworth, J. C., and Swinney, P. A. (2008). An agent-based model of the UK housing market. Technical report, CRESS University of Surrey, http://cress.soc.surrey.ac.uk/housingmarket/ukhm.html.

GRÜNE-YANOFF, T. (2009). The explanatory potential of artificial societies. *Synthese*, 169(3):539-555. [doi:10.1007/s11229-

008-9429-0]

HAMILL, L. and Gilbert, N. (2009). Social circles: a simple structure for agent-based social network models. *Journal of Artificial Societies and Social Simulation*, 12(2):3 http://jasss.soc.surrey.ac.uk/12/2/3.html.

HEDSTRÖM, P. and Swedberg, R. (1998) *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge University Press. [doi:10.1017/CBO9780511663901]

HEDSTRÖM, P. and Ylikoski, P. (2010). Causal mechanisms in the social sciences. *Annual Review of Sociology*, 36:49-67. [doi:10.1146/annurev.soc.012809.102632]

HEMPEL, C. G. (1965). *Aspects of Scientific Explanation and other Essays in the Philosophy of Science*. New York: Free Press.

HEMPEL, C. G. and Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15:135-175. [doi:10.1086/286983]

HUMPHREYS, P. (2002). Computational models. *Philosophy of Science*, 69:1-11. [doi:10.1086/341763]

HUMPHREYS, P. (2009). The philosophical novelty of computer simulation methods. *Synthese*, 169(3):615-626. [doi:10.1007/s11229-008-9435-2]

HUMPHREYS, P. and Bedau, M., eds. (2007). *Emergence: Contemporary Readings in Science and Philosophy*. MIT Press.

KITCHER, P. (1976). Explanation, conjunction and unification. *Journal of Philosophy*, 73: 207-212. [doi:10.2307/2025559]

KÜPPERS, G. and Lenhard, J. (2005). Validation of simulation: Patterns in the social and natural sciences. *Journal of Artificial Societies and Social Simulation*, 8(4)3 http://jasss.soc.surrey.ac.uk/8/4/3.html.

LIPTON, P. (2004). *Inference to the Best Explanation*. New York: Routledge, 2nd edition.

MACHAMER, P., Darden, L., and Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, 67(March):1-25. [doi:10.1086/392759]

MACY, M. and Sato, Y. (2002). Trust, cooperation, and market formation in the U.S. and Japan. *PNAS*, 99: 7214-7220. [doi:10.1073/pnas.082097399]

RABINOW, P. and Sullivan, W. M. (1979). *Interpretative Social Science*. University of Chicago Press.

RESCHER, N. (1957). On Prediction and Explanation. *British Journal for the Philosophy of Science* 8:281-290.

SALMON, W. (1978). Why ask, 'Why?'? An inquiry concerning scientific explanation. In *Proceedings and Addresses of the American Philosophical Association* 51: 683-705. [doi:10.2307/3129654]

SALMON, W. (1990). *Four Decades of Scientific Explanation*. Minneapolis, Minnesota: The University of Minnesota Press.

SALMON, W. (1998). *Causality and Explanation*. Cambridge University Press. [doi:10.1093/0195108647.001.0001]

SCHELLING, T. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, 1:143-186. [doi:10.1080/0022250X.1971.9989794]

TESFATSION, L. (2003). Agent-based computational economics: Modeling economies as complex adaptive systems,. *Information Sciences*, 149(4):262-268. [doi:10.1016/S0020-0255(02)00280-3]

THOMPSON, N. S. and Derr, P. (2009). Contra Epstein, good explanations predict. *Journal of Artificial Societies and Social Simulation*, 12(1):9 http://jasss.soc.surrey.ac.uk/12/1/9.html.

TROITZSCH, K. G. (2009). Not all explanations predict satisfactorily, and not all good predictions explain. *Journal of Artificial Societies and Social Simulation*, 12(1):10 http://jasss.soc.surrey.ac.uk/12/1/10.html.

VAN Baal, P. (2004). *Computer Simulations of Criminal Deterrence*. The Federation Press.

VAN Baal, P. (2008). Realistic Spatial Backcloth is not that Important in *Agent Based Simulation*, chapter 2. Hershey, PA: Idea Group Publishing.

VAN FRAASSEN, B. (1980). The pragmatic theory of explanation. In von Fraassen, B., editor, *The Scientific Image*. Oxford University Press.

WINSBERG, E. (2003). Simulated experiments: Methodology for a virtual world. *Philosophy of Science*, 70:105-125. [doi:10.1086/367872]

YILKOSKI, P. (2001). Understanding Interests and Causal Explanation. PhD thesis, Department of Moral and Social Philosophy, University of Helsinki.