GENERATIVE EXPLANATION AND INDIVIDUALISM

IN AGENT-BASED SIMULATION

Caterina Marchionni* and Petri Ylikoski

Authors' contact details:

Caterina Marchionni

Finnish Centre of Excellence in the Philosophy of the Social Sciences

Department of Political and Economic Studies

P.O. Box 24

00014 University of Helsinki

Finland

Email: caterina.marchionni@helsinki.fi

Petri Ylikoski

Finnish Centre of Excellence in the Philosophy of the Social Sciences

Department of Social Research

P.O. Box 24

00014 University of Helsinki

Finland

Email: petri.ylikoski@helsinki.fi

GENERATIVE EXPLANATION AND INDIVIDUALISM

IN AGENT-BASED SIMULATION

Abstract. Social scientists associate agent-based simulation models with three ideas about explanation: they provide generative explanations, they are models of mechanisms, and they implement methodological individualism. In light of a philosophical account of explanation, we show that these ideas are not necessarily related and offer an account of the explanatory import of agent-based simulation models. We also argue that their bottom-up research strategy should be distinguished from methodological individualism.

1. Introduction

Over the last two decades, agent-based simulations have been increasingly employed throughout the social sciences. However, the methodology of agent-based simulation (ABS henceforth) is not yet sufficiently understood, and its legitimacy and implications are still subject of debate. For example, many economists are wary of simulations because they regard them as inferior vis-à-vis analytical models (Lehtinen and Kuorikoski 2007). In contrast, advocates of ABS methodology praise its ability to overcome the limitations of traditional formal modeling tools (Miller and Page 2007; Macy and Flache 2009).

In this paper our focus is on the explanatory uses of ABS (though of course ABS are used for other purposes as well). In particular, we examine three ideas about explanation that are often associated with the methodology of ABS: it provides explanations that are generative (J. Epstein 2006) and

mechanistic (e.g. Gilbert and Ahrweiler 2009; Hedström 2005; Macy et al. 2011) and it is an implementation of methodological individualism (e.g. Macy and Flache 2009; Neumann 2008; B. Epstein 2012). We argue that, contrary to what it is sometimes believed, these ideas are not necessarily related. With the help of a theory of explanation, we offer an account of how ABS models explain. We show that the ideas of generation and mechanism are sufficient to define the bottom-up research strategy of ABS, but that the latter should not be conflated with the doctrine of methodological individualism.

The structure of the paper is the following. Section 2 briefly introduces the main features of ABS in the social sciences. Section 3 presents our account of how model-based simulations provide explanations. In section 4 we show how the idea of bottom-up explanation is related to ideas of generative and mechanism-based explanation. Section 5 brings up the idea of methodological individualism. Section 6 presents an example of ABS study and shows that some of its key explanatory variables are structural rather than individual properties. Finally, section 7 concludes our argument that it is misleading to regard ABS models as implementations of methodological individualism.

2. Agent-based simulation in social science

In this section we briefly summarize the main features of ABS as they are typically employed in social science. ABS includes a set of agents and a set of rules describing the behavior of those agents. Agents interact with each other and with the environment, and these local interactions often bring about surprising phenomena at the social level. As summarized by Macy and Flache (2009), the agents in ABS models are typically endowed with interesting properties at both the cognitive and social level. Agents are heuristic in that they follow simple behavioral rules that can be interpreted as habits, rituals, routines, norms, and the like. They are also adaptive in that they respond to feedback from their environment through learning and evolution. Agents are autonomous and able to change aspects of their environment to attain their goals. Their autonomy, however, is constrained: agents are assumed

to be interdependent insofar as the local environment to which they react is constituted by the beliefs, goals, and behaviors of other agents. ABS methodology also permits the modeling of heterogeneous agents, namely agents that differ in their beliefs, goals and rules of behavior. Finally, agents can be embedded in networks. All this means that population dynamics are emergent outcomes of local interactions. ABS also allows agents to change their structural locations or break off relations with their neighbors and seek out new relations (Miller and Page 2007, 80-87; Macy and Flache 2009, 247).

Compared to the traditional modeling tools employed in the social sciences, ABS is claimed to have a number of advantages. First, as we have seen, simulations allow one to be flexible about the characteristics of agents and hence break away from the strictures of the assumption of optimizing behavior characteristic of rational choice models. Furthermore, ABS allows modelers to study equilibrium outcomes as well as the dynamics of systems. Finally, unlike standard economic models, where for reasons of tractability the modeler has to work with cases with either one, two or an infinite number of agents, ABS can work with any number of agents and hence can provide more realistic models of social processes (Miller and Page 2007; Macy and Flache 2009; Ylikoski 2014.).

3. Explaining with simulation models

Our account of explanation is realistic in spirit, for it is about tracking the network of dependencies related to the phenomenon of interest. In other words, scientific understanding is constituted by knowledge of dependencies, knowledge that gives us the ability to make correct *what-if* inferences about the phenomenon (Woodward 2003; Ylikoski and Kuorikoski 2010). The point of (explanatory) models is to represent those dependencies in a cognitively salient way, so that they can be used to make what-if inferences about the phenomena under investigation. In ABS models we can call assumptions that pick out explanatorily relevant factors '*explanans* variables' and the outcome to be explained the '*explanandum* variable'.

In this account, mechanisms can be understood as descriptions of the networks of counterfactual dependencies that characterize the system in question. Knowledge about the components of the system, their properties and their organization make it possible to understand why the counterfactual dependencies that characterize the system are the way they do and what their background conditions are (Woodward 2003; Craver 2007; Ylikoski & Kuorikoski 2010). One of the advantages of ABS methodology is that it allows systematic study of how the organization of the components affects the overall behavior of the system.

How can ABS models be used to provide explanations? To answer this question, we distinguish two elements in the process of using ABS for explanation. First, there is *the generative component* of the simulation practice. Here, the simulator sets the rules of interaction for the agents and runs the simulation to produce some macro outcome. If the macro outcome has the characteristics of interest (such as ending with a segregated equilibrium (Schelling 1978) or creating a pattern found in empirical material (J. Epstein 2006)), then we have a proof of possibility. In such cases the simulation demonstrates that the mechanism implemented therein can, in principle, generate the pattern of interest.

While such a proof of possibility is valuable, its contribution to explanation is limited. The simulation does not allow us to identify the conditions under which the mechanism produces the outcome nor which assumptions of the model are responsible for the outcome of interest. In other words, the simulation does not provide us an understanding of how the outcomes systematically depend on the assumptions of the model. Such an understanding is obtained in *the experimental component* of simulation practice.

The experimental component consists of the systematic variation of the simulation assumptions[1]. The point of such 'experimentation' is to see how the assumptions make a difference to the outcome. It

---

[1] Mäki (2005) and Morgan (2003) compare the manipulation of modeling assumptions to experimental manipulation.

provides an understanding of how the outcome of the simulation depends on various assumptions, thus providing knowledge about the relevant networks of dependence. On the basis of this knowledge, the simulator can make what-if inferences about the simulation. As her ability to make such inferences increases, her understanding of the simulation model also increases.

All ABS models include a generative component, but some almost completely lack an experimental component. An example of such a simulation is Kirman and Vriend's simulation of the Marseille fish market (Kirman and Vriend 2001). The stated goal of this model is to explain two stylized facts that characterize the fish market in Marseille, namely the high loyalty of buyers to sellers and the persistence of price dispersion (Kirman and Vriend 2001, 459). These characteristics are puzzling, as the same buyers and sellers meet every day in the same place. In their ABS model there are two kinds of agents, namely buyers and sellers. Sellers decide on the quantities to supply, the prices and whether or not to treat loyal buyers differently. Buyers only decide which sellers to visit and which price to accept. Here, it is assumed that the agents learn through reinforcement. In the words of Kirman and Vriend:

> The model explains both stylized facts, price dispersion and high loyalty. In a coevolutionary process, buyers learn to become loyal as sellers learn to offer higher utility to loyal buyers, while these sellers, in turn, learn to offer higher utility to loyal buyers as they happen to realize higher gross revenues from loyal buyers. (Kirman and Vriend 2001, 459.)

In our view the authors may be overstating the extent to which the model *explains* the stylized facts. This is not to understate the importance of this model and its result. As their simulation can generate high loyalty and price dispersion, it achieves an abstract proof of possibility. Such a proof shows what kind of assumptions could produce the outcome, but not how that occurs or whether those assumptions are the only way to generate the outcome of interest. The model does not provide much insight into the crucial how-questions, as the experimental component is almost completely absent.

6

(The authors only vary one condition, namely the heterogeneity of buyers). Thus, insofar as the authors do not systematically explore how the outcome of interest depends on the details of the model, the full explanatory import of the model remains an open question.

On our account of explanation, to provide a proper explanation of the phenomenon in question, the simulator should show not only that the assumptions made about the agents bring about the observed macro outcome, but also *how* they do so. This is done by spelling out the *mechanism* implemented in the simulation and by showing how the differences in outcomes systematically depend on changes in the assumptions of the simulation. When simulation practice involves an experimental component, the assumptions of the simulation are systematically varied in order to learn the effects of these changes on the outcome of the simulation. This makes it possible to learn more about the network of counterfactual dependencies that characterizes the simulated system. This network of dependencies is precisely what the description of the explanatory mechanism implemented in the simulation amounts to.

Note that there are two reasons for the systematic variation of the assumptions of a simulation. The first is to learn about the systematic dependencies between the *explanans* and the *explanandum* within the specified mechanical configuration. The second is to learn which unrealistic assumptions matter for the model's results, a purpose that is typically achieved by *robustness analysis* (Levins 1966; Wimsatt 1981; Weisberg 2006; Kuorikoski, Lehtinen and Marchionni 2010). Both purposes are equally important for the explanatory use of simulations, but the contribution of robustness analysis to explanation is only indirect. Usually robustness analysis is aimed at investigating the role of those assumptions made to facilitate the tractability of the underlying model or of the simulation itself (as for instance when a city is represented as a checkerboard). Such assumptions are known to be unrealistic but are made because replacing them with realistic assumptions is either difficult or impossible in the first place. In robustness analysis such assumptions are usually replaced with other

unrealistic assumptions (as for instance when an environment is modeled as a grid and then as a torus). The point is to learn whether the change makes a difference to the simulation outcome. If the result is robust with respect to the change, the researcher can infer that the given unrealistic assumption does not play a crucial role in the model and thus it is not an important explanatory variable in the model.

As the purpose of robustness analysis is to identify those assumptions that do not play a central role in the model, it does not matter whether they are realistic. In contrast, explanatory analysis targets those assumptions that do affect the simulation outcomes. Here, it is important that the assumptions *make a difference* within a realistic range of variation. Only these assumptions enable inferences to real-world systems – even though whether a given modification has a reasonable empirical counterpart is an empirical question through and through.

Sometimes modellers do explicitly distinguish between the two kinds of manipulations, but this is not the main criterion for drawing the distinction. In many cases robustness analysis can direct attention to factors that were previously thought to be irrelevant (or vice versa), thus defying the modeller's expectations. In this sense robustness analysis and the variation of assumptions for explanatory purposes are complementary and often concurrent practices. At an abstract level the purpose of both is to explore the properties of both the model and its components in order to learn what depends on what.

More importantly for our purposes, the distinction between the two different aims of variation of a simulation's assumptions provides the criteria for identifying *explanatory* variables in ABS models: we should look for assumptions that are varied by the simulators (in a more or less systematic way), make a difference to the *explanandum* variables of interest and are assumed to have a realistic sociological interpretation. We will illustrate these points in Section 6 with the help of a simulation model of the spread of unpopular norms, which includes manipulations of various of its assumptions.

4. The bottom-up strategy of explanation in ABS

The typical aim of ABS is to provide explanation for macro phenomena such as the emergence and dynamics of system level properties (e.g. distributions of beliefs and behaviors within the population, levels of segregation, networks between the agents). In contrast to more traditional equation-based simulation approaches, ABS does not directly relate macro-variables to each other; rather, it represents a *bottom-up research strategy* that investigates macro phenomena by focusing on agents and their properties, behaviours and interactions (bottom). It shows how what occurs at this level produces, or can produce, the phenomena of interest (up). When building the model, the simulator specifies the rules of behavior for the parts of the system (the agents) and studies how their local interactions generate the relevant macro outcomes. Thus, the simulation requires an explicit consideration of the relevant causal mechanisms. This is in sharp contrast to variable-based simulations, which often operate with macro-variables and do not include explicit thinking about mechanisms. Furthermore, the macro outcomes of ABS models are often surprising, which supports the idea of a bottom-up research strategy. The concept of bottom-up research strategy suffices to make sense of two ideas about explanation in ABS: generation and mechanism. We briefly examine each in turn and then argue that methodological individualism does not capture anything useful that is not captured by the idea of bottom-up explanation.

4.1. Generative explanation

As the purpose of ABS is to obtain macro outcomes from assumptions about agents, their interdependent behaviors and the features of the environment in which they act, it is quite natural then to interpret ABS models in terms of *generative explanation*. Joshua Epstein (2006) gives the strongest formulation of the idea of generative explanation. According to Epstein, growing the macro-level outcome *explanandum* with ABS is a necessary condition for its explanation. This is the origin of the well-known slogan: "*if you didn't grow it, you didn't explain it*" (J. Epstein 2006, 8-10). Epstein's claim

points towards an important feature of ABS: if we are interested in the bottom-up explanations of social phenomena, ABS provides a much-needed tool. However, Epstein's claim is overstated (Ylikoski 2014). First, it is reasonable to assume that explanations of social phenomena have been given before the use of ABS. Second, it is often the case that the result obtained in a given simulation could be generated in a number of alternative ways; thus the simulation only provides a *how-possibly* explanation (Ylikoski 2014). Hence, generative sufficiency is an essential explanatory step, but it does not imply that the empirical outcome to be explained was actually produced in that fashion. Third, simply "growing" the phenomenon of interest is not sufficient for a proper understanding of it. Merely having a simulation that – after a lot of tinkering – produces the right kind of results does not yet yield understanding as it does not enable to make *what if* –inferences beyond that particular simulation configuration. The further challenge is to understand *how* the specified micro-configuration produces the phenomenon and what the background conditions of its production are. Thus, Macy and Flache (2009, 263) are right when they challenge Epstein's slogan by claiming that: "*If you don't know how you grew it, you didn't explain it.*" The idea of generation is necessary for explanatory understanding, but it is not sufficient. We also need to understand the working of the simulation and the systematic dependencies that the experimental component aims to uncover. The idea of mechanism aims to capture precisely these elements of ABS.

4.2 Mechanistic explanation

In the ABS literature, the idea of generative explanation is often associated with the idea of mechanism-based explanation (Hedström and Ylikoski 2010), as suggested for example by the following quotation:

> When we write a set of computational algorithms (the program), formalizing the generative hypotheses the consequences of which are to be studied, what we are doing is hypothesizing a series of generative mechanisms. When we execute the program ... we engender the process

deriving from the set of posited generative mechanisms. With the technical distinction between program "writing," "compilation," and "execution" it becomes clear that a "process" is nothing more than the dynamic aspect of one (or several) mechanism(s): it is what the mechanism can trigger. (Manzo 2007, 5–6.)

One of the premises of the mechanistic approach in social science is that a proper explanation of social phenomena requires an understanding of the causal mechanisms that bring them about. This implies paying attention to the entities of which mechanisms are made (the agents, their properties, actions, and relations) rather than treating them as black boxes. That is, for the explanation of social phenomena it is not sufficient to identify the macro-level changes that produce them. It is crucial to show *how* macro-states affect individuals at a certain point in time and how the actions of those individuals produce new macro-states at a later time (Hedström and Ylikoski 2010). From this perspective, ABS represents an especially useful tool for theoretically oriented social scientists: It allows testing whether certain theoretical ideas are, even in principle, sufficient to generate the phenomenon in question (Epstein 2006; Ylikoski 2014). Furthermore, ABS models can easily be interpreted as representing social interactions: the agents can be interpreted as having goals and beliefs, standing in various social relations to each other, having the resources needed to attain their goals, and also affecting each other's possibilities of action via their behavior. As it should be clear by now, the methodology of ABS, and in particular its experimental component, facilitates the provision of mechanism-based explanations as it enables the systematic study of how the macro phenomenon is generated from assumptions about agents as well as from various non-individualistic assumptions.

5. Is this methodological individualism?

The bottom-up research strategy suffices to capture the potential of ABS as tools for yielding both generative and mechanistic explanations. We have seen that the generative and experimental component of simulation practice enables the simulator to explain macro phenomena by focusing on

agents and their properties, behaviours and interactions and by showing how what occurs at this level

generates or can generate the phenomena of interest. Many authors, however, have suggested that

ABS is an implementation of methodological individualism, or that there is an inherent connection

between ABS and methodological individualism. For example Macy et alii (2011, 252) write that ABS

"modeling is a formal implementation of 'methodological individualism' … ." (see also Flache and Macy

2009; Neumann 2008; Epstein 201.) The ideas of generative mechanisms and methodological

individualism are sometimes packaged together. For example, Keith Sawyer states that "[a]rtificial

societies are mechanistic in the contemporary sense. They are also firmly methodologically

individualist." (Sawyer 2004, 263; see also Sawyer 2003, 340). Similarly, Macy et. al. (2011: 252) write

that ABS modeling "… is a formal implementation of methodological individualism." (see also Flache

and Macy 2009; Neumann 2008; Epstein 2012.) Our goal is to show that there is no inherent (let

alone necessary) connection between ABS, the generative mechanism-based explanations they can

deliver and methodological individualism. In fact, as we will argue, there is much to gain by breaking

down this connection.

What leads some practitioners and commentators to associate ABS with methodological individualism

is quite clear. In any sensible interpretation, methodological individualism would imply something like

a bottom-up research strategy (although, of course, not all methodological individualists are likely to

be excited about the ABS methodology). However, the crucial question is whether the opposite

implication holds true. Here the problem is in discerning what the doctrine of methodological

individualism actually claims. As is well known, methodological individualism has had a variety of

meanings and interpretations over the years (Pettit 1993; Kincaid 1997; Udehn 2001; Hodgson 2007).

Consequently, the term lacks a clear and generally accepted definition. This is a problem for both

advocates and critics of methodological individualism. Since we believe that the debate over the

proper definition of methodological individualism is a distraction from the real methodological issues,

we will not review the many interpretations of methodological individualism available in the literature. Instead, we will work with a definition of methodological individualism that is substantive enough to be non-trivial but that retains the typical connotation of the concept.

Therefore, we take the doctrine of methodological individualism to roughly pertain to the following:

> [MI] Social phenomena can only be explained (or they are best explained) by accounts that only refer to individuals, their properties and their interactions.

There are two aspects to note about this reading of methodological individualism. First, [MI] is a thesis about explanation, not about ontology. Therefore, arguments about the existence of social wholes, structures and such entities vis-à-vis individuals do not directly bear on arguments about explanation of social phenomena. Second, [MI] qualifies as a *strong* version of methodological individualism in that it holds that explanation of social phenomena should appeal only to individuals, their properties and interactions. The corollary of such a view is that non-individual properties are denied non-derivative explanatory status.

Some individualists endorse weaker versions of the thesis, according to which non-individual properties can play a (non-derivative) explanatory role via their effects on individuals. It is legitimate to ask, however, in what sense these more liberal positions are individualistic and whether they represent the same position that many anti-individualists are trying to defend (see Udehn 2001). As said above, we do not want to get entangled in debates about the proper definition of individualism. For our argument, it is sufficient that [MI] is sufficiently close to what is commonly understood as the content of the methodological individualist doctrine.

6. An example: A model of the spread of unpopular norms

Our strategy against the association of ABS with methodological individualism is to argue that if some of the key explanatory variables in ABS models cannot be meaningfully interpreted as 'individualistic', then ABS explanations should not be regarded as implementations of methodological individualism. To

make such an argument, we need a real example of an explanatory ABS model. We have chosen a

simulation model designed by Damon Centola, Robb Willer and Michael Macy (2005) as we think it is

one of the best examples of the use of an ABS model in theoretical sociology and the authors have

elsewhere suggested that there is a close relationship between ABS and methodological individualism

(Macy et. al. 2010).

The model implements the "emperor dilemma" familiar from the Hans Christian Andersen's fable. In

the model, agents must decide whether to comply with and enforce a norm that is supported by a

few fanatics and opposed by the vast majority. The model is used to examine the population level

implications of the use of norm enforcement to falsely signal genuine conviction. The idea is to study

whether a very small fraction of true believers can spark a cascade of conformity and false

enforcement that quickly engulfs a vulnerable population. Thus, the norm does not become enforced

because people are converted to new beliefs; rather, it is because they feel the need to affirm the

sincerity of their (false) conformity.

In the simulation, the population consists of agents who differ in their beliefs and convictions. A small

group of true believers is assumed to have such strong convictions that they always comply with the

norm. When dissatisfied with the level of compliance of others, they may enforce the norm. The

remainder of the population consists of disbelievers who privately oppose the norm, but with less

conviction than that of the true believers. The disbelievers may deviate from the norm or even

pressure others to deviate as well. However, the disbelievers can also be pressured to support the

norm and even to enforce it. At every iteration of the simulation, each agent observes how many of

his neighbors are complying with the norm and how many are deviating. They also observe how many

neighbors are enforcing the compliance and how many are enforcing deviations from the norm. Based

on this information, the agents decide whether to comply or deviate and whether to force others to

behave similarly in the next round.

In the paper, Centola and his associates report the manipulation of three kinds of variable: 1) access to information about the behavior of other agents, 2) the frequency distribution and clustering of true believers, and 3) the network topology (Centola et. al. 2005: 1021). The results of these simulations are surprising: cascades are much easier to achieve than expected. A small group of true believers can bring about a cascade in population where neighborhoods are local; however they are unable to do so in fully connected populations. Moreover, the clustering of true believers turns out to be relevant: a very small cluster of believers can trigger a cascade, while a great number of randomly distributed believers cannot. Finally, when a small number of random ties reduce the overlap between local neighborhoods, cascades are prevented. On the basis of these observations, the authors conclude that unpopular norms thrive on local misrepresentations of the underlying population distribution; that is, cascades are outcomes of a sampling problem. However, the most interesting result is that disbelievers are crucial for the emergence of cascades. Without them cascades do not begin, and if the agents start to convert into true believers, the following of the norm might paradoxically collapse.

Now let us take a closer look at the manipulated variables. From the point of view of our argument, the crucial question is whether they are individual or structural properties. By *structural properties* we mean properties that are attributed to larger scale entities than individuals or if they are attributed to individuals, they presuppose some larger scale entities. These non-individual properties constitute a rather heterogeneous class; what they share in common is the property of being non-individual properties (Ylikoski 2012). The network topology is clearly a structural assumption about the *macro structure of the population*. It is a structural property in the sense that there is no meaningful way to attribute it to an individual; it is always attributed to a larger scale entity. Similarly, the frequency and degree of clustering of true believers (and other agents) is a *population-level attribute* that cannot be applied to individuals. Finally, while the access to information about other agents is attributed to an individual agent, it is more properly understood as a structural assumption about *relations between*

agents. The relations between individual agents and the overall configuration of these relations in the population are population-level attributes that do not apply to individuals.

Thus, all three key variables are rather prototypical non-individual structural macro properties (Ylikoski 2012, 26-33). Furthermore, all three satisfy our suggested criteria for explanatory variables. First, they make a difference to the outcome, as whether or not a cascade is triggered depends on them. Second, they also have a realistic sociological interpretation, as they capture the degree to which agents can obtain an accurate picture of how widespread genuine belief in a given norm is.

Explanatory structural variables like these are not unique to this particular case. It is quite common to find ABS models that focus on variables such as the composition of the population, contacts between agents, and agents' freedom of movement (e.g. Flache and Macy 2011, Centola and Macy 2007). This is the case even for the relatively simple segregation models inspired by Thomas Schelling's work (e.g. Benard and Willer 2007; Bruch and Mare 2006; Clark and Fossett 2008; Fossett 2006; also see Ylikoski 2014). One of the attractions of ABS modeling is precisely that such assumptions can be systematically manipulated together with assumptions about individuals. For example, the effects of changes of the size of the neighborhood with which agents interact or of the connections that agents have beyond their immediate neighborhood can be studied, for example, by a random rewiring of the links between agents (e.g. Centola and Macy 2007). The agents' freedom of movement brings in another set of structural assumptions, such as the relative size of the available empty spaces and the rules for movement across them. The 'experimental' manipulation of structural and individual assumptions makes it possible for the simulator to investigate the network of dependencies (i.e. the mechanism) that characterizes the phenomenon to be explained. What counts as part of mechanism depends on what makes a difference to the macro-phenomenon, regardless of whether the variables are structural or individualistic.

7. Breaking the association between methodological individualism and ABS

As the model of Centola and colleagues demonstrates, some of the crucial explanatory variables in ABS are structural. This constitutes a prima facie case against the association of ABS with methodological individualism. The burden of proof then shifts to those who wish to make the association: they should show that those explanatory variables are either not explanatory, or that they can be credibly interpreted as individualistic. The discussion above makes clear that, denying their explanatory relevance is not a viable strategy; reinterpreting them as individualistic properties does not look a very promising strategy either. In this case the challenge is to find a *credible* formulation of individualistic properties that could still accommodate the relevant explanatory variables. The definition of individualistically acceptable properties should not be allowed to trivialize the debate between traditional individualists and their critics. To make this argument more compelling, we offer some substantial arguments against the strong version of methodological individualism and some points of caution against associating ABS with methodological individualism that apply to both its strong and weak formulations.

Our first substantial argument is based on the observation that ABS methodology does not dictate how the agents are to be interpreted: while in many applications it is natural to interpret agents as individuals, they can also be households, groups, organizations, or even intrapersonal cognitive processes. The interpretation of the simulation only requires that it makes sense to ascribe to the agents the behavior defined by the rules. Consequently, even within the same simulation, agents can be of various kinds. Thus, in this respect ABS is not by itself individualistic, and restricting its use to individualistically acceptable applications (that is, applications in which agents are to be interpreted as individuals) would require strong arguments. Whether such arguments can be provided is an open question in which advocates of methodological individualism carry the burden of proof.

Another strategy for the individualist is to claim that the structural assumptions we have identified as

doing explanatory work do not pose a problem because structural properties can be shown to result from purely individualistic processes. For example, one might claim that network structures are generated by certain kinds of individual preferences in interaction, thus making them individualistically acceptable. Our replies to this objection constitute the second argument against associating ABS with methodological individualism. First, within a given ABS variables such as network structures are both structural *and* explanatory. That these variables could be themselves explained in individualist terms does not imply that their explanatory relevance vanishes. This kind of regress argument does not work in the case of explanation: the explanatory status of a variable does not depend on whether it can itself be explained (Ylikoski 2012, 40-41). Second, the individualist would still have to demonstrate that it is, in fact, possible to give a purely "individualistic" explanation for these kinds of variables. It is difficult to see how this might happen (Kincaid 1996); nevertheless, the burden of proof is again with the individualist.

Even if the individualist were to stretch her idea of methodological individualism to cover both the cases presented above, what would be the benefit of such a dialectical victory? Probably none.

Our first point of caution is the observation that the track record of the debate around methodological individualism in resolving real micro-macro problems in the social sciences is controversial. It can be argued that such problems can be discussed more fruitfully without getting bogged down by issues related to the proper definition of "individual properties" and "individual level," or by attempts to clarify different meanings of methodological individualism.[2] The notion of bottom-up strategy of explanation we have articulated above suffices to capture both the idea of generation and the virtue of understanding mechanisms, without implying any commitments to principles like [MI].

Moreover, the association of ABS with methodological individualism can make some social scientists

---

[2] Ylikoski (2012) offers an account of macro-micro relations that dispenses with many of the problems that afflict the individualism-holism debate.

less receptive to the productive possibilities of ABS methodology. Considering that the narrow and negative understanding of "methodological individualism" is common within the social scientific community, an overly close association between ABS and methodological individualism will only turn some potentially interested people away from ABS.

Finally, there is the danger that a conceptualization of micro-macro problems in terms of methodological individualism may lead to biased strategies in ABS research (Wimsatt 1983). For example, understanding the agents in ABS exclusively as individuals might make researchers blind to large-scale structural factors that have significant explanatory import and that can be effectively modeled with ABS methodology. Similarly, thinking of properties like network topologies as explanatorily inert may lead to researchers systematically choosing to make simplifications and idealizations at this level rather than at the level of the individuals. In turn, this could lead to a situation in which more realistic accounts of the networks in which individuals are embedded are seldom tried out. Both kinds of bias would be unfortunate for the development of ABS research because one of its advantages is precisely that it allows social scientists to overcome (some of) the traditional limitations of social scientific model-building and theorizing.

It is important to note that our argument against methodological individualism does not in any way imply that the traditional opponents of individualism are right. Our point is that the association of ABS with methodological individualism does nothing to advance our understanding of how ABS models explain, while at the same time it may turn out to be an obstacle to their fruitful use. It may be true that agent-based simulations exemplify many epistemic virtues, such as an explicit consideration of mechanisms and a focus on agents, which traditional methodological individualists have found important. However, these virtues are independent of methodological individualism. Thus, our suggestion is to ignore the individualism-holism debate and focus instead on how the bottom-up research strategy that characterizes ABS can offer new ways of studying micro-macro relations

(Ylikoski 2012).

8. Conclusion

Philosophical reflections on social simulations have been scarce compared to both general accounts of the epistemology of simulations and specific analyses of simulations in the natural sciences (see, however, B. Epstein 2012; Grüne-Yanoff and Weirich 2010.) In this paper we have offered an account of the way in which ABS is used for explanatory purposes in social science. The bottom-up research strategy embodied in typical ABS delivers generative explanations: it generates the macro phenomenon to be explained by appeal to the actions and interactions of the agents. The bottom-up research strategy of ABS can also yield mechanism-based explanations: it tracks the (possible) network of dependencies behind the phenomenon to be explained. However, we argue that the idea that ABS is an implementation of methodological individualism is misleading because structural assumptions often play an irreducible explanatory role. Finally, we have offered practical (rather than conceptual) arguments for resisting the association of ABS with methodological individualism, even when the latter is interpreted to accommodate the explanatory role of non-individualistic assumptions.

References

Benard, Stephen and Robb Willer. 2007. "A Wealth and status-based model of residential segregation." *Journal of Mathematical Sociology* 31(2): 149-174.

Bruch, Elizabeth and Robert Mare. 2006. "Neighborhood choice and neighborhood change." *American Journal of Sociology* 112(3): 667-709.

Centola, Damon and Michael Macy. 2007. "Complex contagions and the weakness of long ties."

American Journal of Sociology 113(3): 702-34.

Centola, Damon, Robb Willer and Michael Macy. 2005. "The Emperor's Dilemma: A Computational Model of Self-Enforcing Norms." *American Journal of Sociology* 110(4): 1009–1040.

Clark, William A. V. and Mark Fossett. 2008. "Understanding the social context of the Schelling segregation model." *PNAS* 105 (11): 4109-4114.

Craver, Carl F. 2007. *Explaining the Brain: mechanisms and the mosaic unity of neuroscience*. Oxford: Clarendon Press.

Epstein, Brian. 2012. "Agent-based modeling and the fallacies of individualism." In *Models, Simulations and Representations*, eds. P. Humphreys and C. Imbert, 115-144. New York: Routledge.

Epstein, Joshua M. 2006. *Generative Social Science. Studies in Agent-Based Computational Modeling*. Princeton: Princeton University Press.

Flache, Andreas and Michael W. Macy. 2011. "Small worlds and cultural polarization." *The Journal of Mathematical Sociology* 35(1): 146-176.

Fossett, Mark. 2006. "Ethnic preferences, social distance dynamics, and residential segregation: theoretical explorations using simulation analysis*." Journal of Mathematical Sociology* 30(3-4): 185-273.

Gilbert, Nigel and Petra Ahrweiler. 2009. "The epistemologies of social simulation research." In *EPOS 2006*, LNAI 5466, ed. F. Squazzoni, 12-28. Berlin: Springer-Verlag.

Grüne-Yanoff, Till and Paul Weirich. 2010. "The philosophy and epistemology of simulation: A review." *Simulation and Gaming* 41(1): 20-50.

Hedström, Peter. 2005. *Dissecting the Social: On the Principles of Analytical Sociology.* Cambridge: Cambridge University Press.

Hedström, Peter and Petri Ylikoski. 2010. "Causal Mechanisms in the Social Sciences." *Annual Review of Sociology* 36: 49-67.

Hodgson, Geoffrey M. 2007. "Meanings of methodological individualism." *Journal of Economic Methodology* 14(2): 211–226.

Kincaid, Harold. 1996. *Philosophical Foundations of the Social Sciences*. Cambridge University Press.

Kincaid, Harold. 1997. *Individualism and the Unity of Science. Essays on Reduction, Explanation, and the Special Sciences*. Lanham: Rowman & Littlefield.

Kirman Alan P. and Nicolaas J. Vriend. 2001. "Evolving market structure: An ACE model of price dispersion and loyalty." *Journal of Economic Dynamics and Control* 25: 459-502.

Kuorikoski, Jaakko, Aki Lehtinen and Caterina Marchionni. 2010. "Economic modelling as robustness analysis." *British Journal for the Philosophy of Science* 61(3): 541-567.

Lehtinen, Aki and Jaakko Kuorikoski. 2007. "Computing the perfect model: why do economists shun simulation?" *Philosophy of Science* 74(3): 304-329.

Levins, Richard. 1966. "The strategy of model building in population biology." *American Scientist* 54(4): 421-431.

Macy, Michael W. and Andreas Flache. 2009. "Social dynamics from the bottom up: agent-based models of social interaction." In *The Oxford Handbook of Analytical Sociology*, eds. P. Hedström and P. Bearman, 245-268. Oxford University Press.

Macy, Michael. W., Damon Centola, Andreas Flache, Arnout ven de Rijt and Robb Willer. 2011. "Social mechanisms and generative explanations: computational models with double agents." In *Analytical Sociology and Social Mechanisms*, ed. P. Demeulenaere, 250-265. Cambridge: Cambridge University Press.

Manzo, Gianluca. 2007. "Comment on Andrew Abbott." *Sociologica* 2: 1–8.

Miller, John H. and Scott E. Page. 2007. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life.* Princeton: Princeton University Press.

Morgan, Mary S. 2003. "Experiments without material intervention: model experiments, virtual experiments and virtually experiments." In *The Philosophy of Scientific Experimentation*, ed. Hans Radder, 216-35. Pittsburgh: University of Pittsburgh Press.

Mäki, Uskali. 2005. "Models are experiments, experiments are models." *Journal of Economic Methodology* 12(2): 303-15.

Neumann, Martin. 2008. "Homo socionicus: a case study of simulation models of norms." *Journal of Artificial Societies and Social Simulation* 11(4): article no. 6. <http://jasss.soc.surrey.ac.uk/11/4/6.html>

Pettit, Philip. 1993. *The Common Mind. An Essay on Psychology, Society, and Politics*. Oxford University Press.

Sawyer, R. Keith. 2003. "Artificial societies: Multi agent systems and the micromacro link in sociological theory." *Sociological Methods and Research* 31(3): 325-363. Reprinted in *Computational Social Science*, ed. N. Gilbert 2010. Sage.

Sawyer, R. Keith. 2004. "The mechanisms of emergence." *Philosophy of Social Sciences* 34(2): 260-282.

Schelling, Thomas C. 1978. *Micromotives and Macrobehavior*. London and New York: W. W. Norton.

Udehn, Lars. 2001. *Methodological individualism: Background, history, and meaning*. London: Routledge.

Udehn, Lars. 2002. "The Changing Face of Methodological Individualism." *Annual Review of Sociology* 28: 479-507.

Weisberg, Michael. 2006. "Robustness analysis." *Philosophy of Science* 73(5): 730-42.

Wimsatt, William C. 1981. "Robustness, reliability and overdetermination." In *Scientific Inquiry and the Social Sciences*, eds. M.B. Brewer and B.E. Collins, 124-63. San Francisco: Jossey Bass.

Wimsatt, William C. 1983. "Reductionistic research strategies and their biases in the units of selection controversy." In *Conceptual Issues in Ecology*, ed. E. Saarinen, 155-201. Dordrecht: D. Reidel Publishing Company.

Woodward, James. 2003. *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.

Ylikoski, Petri. 2012. "Micro, Macro, and Mechanisms", in *The Oxford Handbook of Philosophy of the Social Sciences*, ed. H. Kincaid, 21-45. Oxford University Press.

Ylikoski, Petri. 2014. "Agent-Based Simulation and Sociological Understanding", forthcoming in *Perspectives on Science*.

Ylikoski, Petri and Jaakko Kuorikoski. 2010. "Dissecting Explanatory Power." *Philosophical Studies* 148: 201–219.