Sensemaking of causality in agent-based models

Patrycja Antosz, Timo Szczepanska, Loes Bouman, J. Gareth Polhill & Wander Jager

Published online: 25 Mar 2022.

Submit your article to this journal ⬈

Article views: 581

View related articles ⬈

View Crossmark data ⬈

Routledge
Taylor & Francis Group

Sensemaking of causality in agent-based models

Patrycja Antosz[a], Timo Szczepanska[b], Loes Bouman[c], J. Gareth Polhill[d] and Wander Jager[c]

[a]Department of Health and Social Sciences, NORCE, Kristiansand, Norway; [b]Norwegian College of Fishery Science, UiT the Arctic University of Norway, Tromsø, Norway; [c]University College Groningen, University of Groningen, Groningen, The Netherlands; [d]Information and Computational Sciences, The James Hutton Institute, Aberdeen, UK

ABSTRACT

Even though agent-based modelling is seen as committing to a mechanistic, generative type of causation, the methodology allows for representing many other types of causal explanations. Agent-based models are capable of *integrating* diverse causal relationships into coherent causal mechanisms. They mirror the crucial, multi-level component of emergent phenomena and recognize the important role of single-level causes without limiting the scope of the offered explana- tion. Implementing various types of causal relationships to complement the generative causation offers insight into *how* a multi-level phenomenon happens and allows for building more complete causal explanations. The capacity to work with multiple approaches to causality is crucial when tackling the complex problems of the modern world.

Introduction

The majority of present-day societal challenges (e.g. adapting to climate change, protecting biodiversity, living in greener cities, managing a pandemic and ensuring soil health and food security) require not only novel technological breakthroughs, but also a deeper understanding of multilevel social phenomena. Agent-based modelling (ABM) is widely seen as an inter- and/or transdisciplinary methodology that is able to provide insight into such phenomena through representing the way multilevel, complex systems operate. In this article we argue that, even though ABMs are primarily perceived as committing to a mechanistic, generative type of causation, the methodology allows for the integration of other types of causation in a single research endeavour.

The next section briefly describes the sociological approaches to investigating multilevel phenomena, and concludes by highlighting the importance of mechanism-based explanations. Section 2 summarizes the stance that ABMs provide mechanistic, generative explanations of multilevel social phenomena. In section 3, we substantiate and augment that claim by showing how ABMs have successfully represented various types of causal mechanisms in the past. We conclude by making a normative claim: that ABM not only can, but should, use all available information about causal mechanisms in its attempt to build better models of multilevel social phenomena.

1. Sociological approaches to studying multilevel social phenomena

Investigating multilevel social phenomena has historically been the domain of sociology. The linkages between micro- and macro-levels have fascinated sociologists for decades and are still hotly debated in sociological theory. Even though *sociological theory has not resolved its micro-*

CONTACT Patrycja Antosz ✉ paan@norceresearch.no 🖅 Department of Health and Social Sciences, NORCE, Kristiansand, Norway.

macro divide any more than other sciences (…) sociological theorists seem rather more obsessed with the problem (Turner, 2001, p. 6). Multiple theoretical camps in western sociology re-formulated the micro-macro distinction into the division between agency and structure, and began arguing which is more important. *Micro-chauvinists* argue for the primacy of agency – it is action that gives rise to macro-level forces. *Macro-chauvinists* emphasize the constraining force of structure and culture that make individual actions predictable. Unsurprisingly, while there were polarizing views on which preceded the other, there was also an attempt to establish a middle ground (e.g. Archer, 1982; Barnes, 2001; Giddens, 1984). In the 1980s, the agency-structure primacy discussion settled with a compromise, at least for a brief period.

Over last two decades, a new and promising direction for investigating multilevel phenomena materialized. Mechanism-based explanations seized the social scientific spotlight (Hedström, 2005; Hedström & Ylikoski, 2010; Pozzoni & Kaidesoja, 2021). They did not explicitly commit to a micro-chauvinistic stance as their advocates do not use terms such as agency or structure. Nonetheless, it is hard to overlook the bottom-up emphasis of the approach, according to which a sufficient explanation has to uncover a mechanism, the *cogs and wheels* underlying the investigated phenomenon (Hedström & Ylikoski, 2010, p. 50).[1] In a functionalist interpretation, a mechanism is a *structure that performs a function as a consequence of its component parts, component operations, and organization* (Bechtel & Abrahamsen, 2005, p. 423). The focus on mechanism-based explanations in social sciences follows the rise of analytical sociology, a school of thought advocating the rejection of black-box explanations such as congruence-law or statistical explanations. According to analytical sociology, a sufficient explanation should pay attention to the process details of a social phenomenon to underpin how explanans and explanandum are linked together (Hedström, 2005). This line of thinking is rooted in structural individualism according to which a*ll social facts, their structure and change, are in principle explicable in terms of individuals, their properties, actions, and relations to one another* (Hedström & Ylikoski, 2010, p. 60). Thus, a mechanistic explanation requires the commitment to structural individualism – the mechanism must be composed of individuals, their properties, actions, and relations to one another (Pérez-González, 2020). Thereby, the mechanism-based approach describes a dynamic causal process that generates the effect of interest (Hedström & Manzo, 2015), instead of providing an exhaustive account of all details of the full causal story (Elsenbroich, 2012). Agent-based modelling, a methodology aligned with the principles of structural individualism, was an important contributor to the rise of analytical sociology, as it offers mechanistic explanations of multilevel social phenomena.

2. ABMs as generative explanatory models of multilevel social phenomena[2]

Scientific explanations are answers to certain kinds of 'why' questions (Salmon & Salmon, 1979, p. 62). It is widely accepted that explaining a phenomenon means providing a causal account for it.[3] Before we outline how agent-based models represent a specific, well-defined approach to causality, a short and more general note displaying the plurality of existing approaches is given. Among others, theories and/or models of causal inference and explanation include regularity theories (Hume, 1739 [1978]), the deductive-nomological model (Hempel & Oppenheim, 1948), the inductive-statistical model (Hempel, 1965), the statistical relevance model (Salmon, 1971), counterfactual theories (Lewis, 1973), the INUS model (Mackie, 1974) and mechanistic theories (elaborated on below). These competing approaches propose different ideas about what can be defined as a single and/or joint cause of a phenomenon and provide diverse structures for how explanation in science should be applied.

The ABM approach to explanation largely aligns with mechanistic complex systems theory.[4] In the version of the latter outlined by Machamer et al.'s (2000),[5] dualist mechanisms are composed of entities and activities (i.e. the dualism of entities and activities). While entities engage in change, it is the activities that are the producers of change. Therefore, a mechanism is not seen merely as a mechanical (push-pull) system, but rather as *entities and activities organized such that they are*

productive of regular changes from start or set-up to finish or termination conditions (Machamer et al., 2000, p. 3). The organization of entities and activities matters in the way they produce the phenomenon of interest: entities must have specific properties and/or be appropriately located, and the activities must have a temporal order, rate, and/or duration. One of the central features of agent-based modelling is that the code is explicit about scheduling of the processes and the roles of various types of agents: at any given moment it is traceable which entity is undertaking what activity. Moreover, the ABM community strongly supports clarity in reporting a model's scheduling (e.g. by using flowcharts (Grimm et al., 2006) or pseudo-code (Grimm et al., 2010)).

In a mechanistic manner, ABMs explain macro-level phenomena by explicitly showing how they were generated through micro-level interactions: *The fundamental social structures and group behaviours emerge from the interaction of individual agents operating on artificial environments under rules that place only limited bounds on each agent's information and computational capacity* (Epstein & Axtell, 1996, p. 6). Early models supported the idea that no central control was needed to achieve a macro-level outcome, and simulating processes of emergence became a key characteristic of the agent-based modelling methodology. A widely cited example is the segregation model of Schelling (1971) and Sakoda (1971), which demonstrated how a fully segregated society could emerge despite only a slight preference of individuals to avoid becoming a minority in a neighbourhood. Another popular example is the flocking of birds, which mesmerised many people and provoked them to speculate about a *group intellect*. As it turns out, complex macro-level patterns could be easily grown from a few simple rules guiding individual bird behaviour (Reynolds, 1987).

There is a general consensus that emergence is by definition a feature of any ABM.[6] It is how we teach agent-based modelling and how we usually build models (e.g. Macal & North, 2010). Epstein (2006) coined the term *generative social science* to address the capability of ABMs to understand complex social phenomena by *growing* them in computer simulations. The idea, originating in the Santa Fe Institute, has dominated thinking about micro-macro linkages among agent-based modellers, and the perception of agent-based modelling among other scholars. The methodology is widely thought of as *a computational implementation of 'methodological individualism,' the search for the microfoundations of social life in the actions of intentional agents* (Macy & Flache, 2009, p. 245; but also e.g. O'Sullivan & Haklay, 2000). As a result, historically, agent-based modelling unjustly acquired a reputation of being a micro-chauvinistic approach, which is a rare achievement for a methodology that was not founded on any particular such theory. To this day, the field is dominated by models that strictly follow methodological individualism and give primacy to human agency. However, the mechanistic generative causation leading to emergence is not the *only* type of causal mechanism that agent-based modelling is capable of representing. Due to its multilevel nature, the methodology can also investigate how the macro-level phenomena impact the behaviour of individuals (often referred to as 'downward causation'). The beginning of the 21st century marked the start of a meaningful and noteworthy debate that recognizes the potential of ABMs to include downward causation and represent bi-directional causality (Conte & Castelfranchi, 1995; Conte et al., 2001; Sawyer, 2000). As a result of the debate, representing bi-directional causality was recommended as a standard guiding future work. For example, Conte and Paolucci (2014) argue for including bi-directional causality when representing social phenomena, and Filatova et al. (2013) when representing socio-ecological systems. Subsequently, examples of ABMs that explicitly implemented multi-level bi-directional causality materialized. Gilbert (2002) modified the Schelling/Sakoda model by adding a neighbourhood-level crime rate, that partially determined the price of housing in the neighbourhood, and restricted the new location moved to by the ratio of old to new property value. Implementing the emergent macro-level structural restriction of individual actions replicated the clustering pattern found in the original study. Conte et al. (2013) provided further examples of both simple and complex (i.e. second-order emergence and immergence) causal loops.

Similar feedback mechanisms and dynamic causal loops are distinctive features of complex social systems mimicked in ABMs (Conte & Paolucci, 2014), and have been previously recognized in sociological theory (e.g. Archer, 1982; Barnes, 2001; Giddens, 1984).

Examples in the last paragraph clearly indicate that the emergence-exclusivity of ABMs has already been contested. Scholars in the past focused mainly on the causal links between the levels, particularly acknowledging macro-micro, and bi-directional causal influences. The next section will further challenge the view that agent-based modelling is exclusively committed to a single account of causality i.e. to building mechanistic, generative explanations of emergent phenomena. We highlight the various within-level causal relationships that can be integrated in a complex mechanism represented by an ABM.

3. ABM as an integrative platform for causal relationships

> Every nut, bolt and washer in a laboratory instrument is the carrier of epistemological assumptions. Every aspect of design, data collection and analysis in social research bears a commitment to a particular model of social explanation.
>
> (Pawson & Tilley, 1996), p. 574

ABMs are capable of representing coherent and relatively complex causal mechanisms by serving as integrative platforms for causal assumptions adhering to various theories of causal explanation, not only to the mechanistic generative emergence. We begin this section by enumerating diverse interpretations of causal relationships that, alongside generative emergence, are popular in social sciences. We present examples of how causal relationships are expressed and point to some of the epistemological assumptions that may be underlying these expressions. Subsequently, to substantiate the claim that ABMs can serve as integrative platforms for diverse causal assumptions, we provide three examples of ABMs that successfully combined various modes of causation into coherent mechanisms without sacrificing the generative identity of the methodology.

The approach to causal inference taken by a given researcher is reflected in their research questions, and the subsequent choice of research design[7] (incl. the implemented research and data collection methods, and techniques to analyse data). The way we ask questions about reality determines the way we answer. The explanations scientists give are indicative of what they perceive as causes and effects, and as the appropriate manner to investigate causal relationships. However, the way that the explanation is formulated is determined *not only* by the research question (and the broadly-defined research design that followed), but also by the scientific background of the scholar, and the dominant idiosyncrasies of their disciplines, especially with respect to the preferred (formal) language to express casual relationships. For example,[8] physical sciences often describe causal relationships in the inorganic world by outlining deterministic formulas. The formulas often follow a deductive-nomological model of explanation (Hempel & Oppenheim, 1948), where the explanandum follows logically from the premises in the explanans (i.e. deductive) and one of the premises of the explanans is a law of nature (i.e. nomological). On the other hand, quantitative approaches in social sciences (e.g. methods and techniques such as experiments and survey questionnaires) usually analytically express social phenomena in a probabilistic manner (e.g. with the use of analytical techniques based on regression; Russo, 2009). There is a number of theoretical approaches to causality that a regression equation can follow. From Hempel's (1965) inductive-statistical explanations where the effect follows a cause(s), given a statistical law (i.e. with high probability, rather than universally), through Salmon's (1971) statistical relevance model where causes should *make a difference* to the occurrence or non-occurrence of the explained effect, rather than determine it with high probability, to Mackie's (1974) INUS model where a cause is an Insufficient but Nonredundant (necessary) part of a condition that is Unnecessary but Sufficient (e.g. a more complex, multivariate regression model with an interaction).[9] Experimental and quasi-

experimental research designs can easily use causal effect or the counterfactual models of causal inference. In a nutshell, in counterfactual approaches the cause is *something that makes a difference, and the difference it makes must be a difference from what would have happened without it* (Lewis, 1986, pp. 160–161). The Neuman-Rubin causal model (Rubin, 1974) defines the causal effect as the difference in the effect when a cause is present, compared to a situation when the cause is absent, and in its extension by Holland (1986, also called the Neuman-Rubin-Holland model), elaborately deals with the fundamental problem of causal inference, namely the issue that direct observation of the causal effect is impossible (at the same point in time the actual state and the counterfactual state cannot not exist together e.g. John either took the medicine or did not take it). Mohr (1982) contrasts such *variance-theory models of explanation* present in quantitative social sciences with *causal processes* (Salmon, 1984) by which some events influence other events. The latter approach has been implemented in qualitative methods (e.g. case studies, in-depth interviews or focus group interviews),[10] although it is disputable if its author would agree with such implementations (Salmon, 1997). Furthermore, comparative studies make use of multiple conjunctural causality (Ragin, 1987). Similar to the INUS model, conjunctural approaches investigate combinations of causes rather than single causal influences, and employ Mill's (1843 [1967]) method of agreement and indirect method of differences to identify conjunctions of cases jointly sufficient for producing a given outcome.

ABMs explicitly define a *causal mechanism* responsible for eliciting the phenomenon under investigation. Micro-level interactions between agents responsible for eliciting the emergent macro-level phenomenon are the core part of that causal mechanism. This way, ABMs adhere to the principles of generative causation. However, ABMs can, and often do, incorporate other types of causation. This happens because elements of the mechanism (Bechtel & Abrahamsen, 2005), such as component parts (e.g. heterogenous agents of different types/breeds), component operations (e.g. actions of agents), and their organization (e.g. causal relationships between components) can be constructed with the use of information from other methods and techniques, even if those approach causality in different ways. We believe that the way causal relationships are expressed by non-modellers is crucial for this discussion, as it constitutes the foundation for the ABM inputs. Modellers most often do not have the opportunity to acquire a first-hand in-depth understanding of the causal relationships of a target system they want to include in the ABM. Consequently, in the quest to make sense of the represented reality and ascribe meaning to the models, they rely on what they manage to find in existing written accounts (empirical or theoretical). As the target system mimicked in the ABM is usually complex, to assure a sufficient level of resemblance modellers utilize all the causal assumptions they can get their hands on.[11] This process forces them to interpret the various languages causal relationship were formulated in. To back our claim with evidence, we present three examples of non-generative causal assumptions implemented in agent-based models: deterministic equations, probabilistic relationships and causal processes.

Example 1: deterministic equations in agent-based models
Classical physics aspires to provide us mortals with the Laws of Nature (well-established, empirically verified principles expressed in a form of equations e.g. Newton's law of gravitation, or his three laws of motion), which can be used to explain a wide range of phenomena. The existence of deterministic, universal laws feels warm and cosy, as it decreases uncertainties and offers psychological comfort. The determinism in classical physics is based on finding an existing, unique solution in the theory of ordinary differential equations (Clark, 1990). Using the Laws of Nature allows for building the covering-law models (Hempel & Oppenheim, 1948), where explanation of a phenomenon is given by a logical conclusion of a deductive argument with law statements and descriptions of initial conditions as premises. Recently, Polhill et al. (2021) show that there might be potential for building law-like principles also in the social world. Assuming explanation-prediction symmetry, the authors successfully demonstrate that neither complexity nor wickedness make finding a solution in a deterministic *isolated* complex system intractable, however endogenous

ontological novelty in wicked systems renders prediction futile beyond the immediately short term. With less optimism, Elsenbroich (2012) argues that covering-law explanations (or, to be precise, *statistical explanations based on the covering law approach*) have limited usability in agent-based models that investigate underlying mechanisms and provide explanations of social phenomena, although they *can be used as evidence for causal associations* (ibid.). They can, and they indeed were. The examples in the ABM literature are plentiful. Here, we point to Marilleau et al. (2018), who simulated the spreading of a montane water vole (*Arvicola scherman*) population in the Haute-Romanche valley (France) by coupling ABM and EBM in a single multiscale model. Agent-based modelling was chosen due to its capacity to represent spatial heterogeneity (i.e. variation in topographic slope that restricts vole movements) and individual micro-scale behaviours (i.e. vole movements determining which agents can interact together and produce offspring). Equation-based modelling offered the analytical solution in the form of a logistic, age-structured population growth model. The submodel expressed how the change (growth) of the young vole population in each cell was determined by the number of voles inside the cell and the carrying capacity of that cell. Reflecting on future work, authors emphasized that other equations (e.g. representing a weather component or a more accurate predation model) can be added to the current model, although a significant challenge of maintaining coherence while increasing complexity would occur: *In this case, the model would include even more scales and would call for a specific organization of calculations and links between different scales* (ibid., p. 40).

Example 2: probabilistic relationships in agent-based models

Combining games with ABMs is still a relatively rare practice (for a review, see Szczepanska et al., in press in this special section of IJSRM). While games were initially applied in participatory modelling as a component of the model design phase, they can also be found as an integral part of an experimental setting. Researchers use experiments to study causal relationships by (1) manipulating the cause and observing the effects afterwards, (2) examining if variation in the cause is linked to variation in effects, and (3) reducing the plausibility of other explanations for the effects (Shadish et al., 2002). The examples of implementing probabilistic relationships in ABMs we give you here come from two studies that combined all three: games, experiments and agent-based modelling. Structural similarities between games and agent-based models allow researchers to operate in two counterpart realities: the game interface of the experiment and the agent-based computer meta-model. Observations from the gameplay are used twofold: (1) to determine probabilities of certain actions that are subsequently used in ABM calibration, and (2) to provide data for ABM validation that is achieved by comparing the simulations with experimental results. The ABM simulations are subsequently used to scale up the game/experimental results by expanding the parameter space. Bhattacharya et al. (2019) investigate the rationality of the decision-making processes of individuals located in social networks. In an online game, players of different colours communicate over their network links and exchange their locations (node positions). The goal is to switch network locations with other players to achieve global colour-clustering. Analysing experimental data with a logistic regression, researchers examine to what extent a player's ratio of sent requests to no request depends on factors like cluster size of the requesting player, cluster size of the requested neighbour, etc. (for a more comprehensive description, see the original paper). In what is labelled as a practice of *probability matching*, authors later use the probabilities of sending a request in various contexts (as defined by the independent variables of the logistic regression) to calibrate the decision-making of the agents. Once evaluated by fitting to the experimental data, the parameters in the ABM are used in numerical simulations as the agent behaviour baseline. Authors then scan the parameter space (by varying agent strategies) and compare the effectiveness of other possible strategies with the strategy employed by humans. Cedeno-Mieles et al. (2020) applied a very similar method combination to investigate how collective identity develops in groups. They also used a (multilevel) logistic regression to determine probabilities of possible actions taken by the

players. However, by implementing a set of independent variables as a vector, the authors used a combination of probabilistic and configurational causality – an idea that an effect is caused by a specific combination of multiple factors that have to come together rather than by individual independent variables.

Example 3: causal processes in agent-based models

Process theory, as a form of causal explanation, is rooted in a realist approach to causation (Maxwell, 2004). Conducting an in-depth study of one or few cases, engaging with a small sample of individuals, or examining qualitative secondary data (e.g. documents, movies, photographs), researchers are able to recreate chronological and contextual connections between events (i.e. causes and effects). Antosz et al. (2020) implemented the approach to the phenomenon of shirking – voluntarily working less than expected by the superior in an organization. After a failed approach to ground the agent-based model in game theory (Antosz & Verhagen, 2020), the authors turned to qualitative, processual approaches for help. A review of the literature provided a picture of factors that play a role in the amount of employee shirking. However, the question of: *how does shirking actually happen?* remained open. To fill in the gaps about the general structure of the work process individual in-depth interviews with employees and managers were carried out. Qualitative information drew attention to processes that have not yet been described in the literature: the crucial role of (1) managerial expectations, against which employee performance is measured, and (2) unexpected events that might either cause delays or speed up the task execution. Acquired information allowed for defining agent types with their role-specific actions, and served as guidance for model scheduling: starting with creating tasks of various complexity, these tasks are further allocated to employees according to manager's imperfect knowledge about employees and tasks, in the last step employees complete the tasks and inform the managers. Moreover, information provided clear rules for role-specific actions. For example, managers only allocated tasks to available employees, and task difficulty corresponded with employee competence (both as perceived by the manager). Last but not least, processes that respondents described allowed for identifying a crucial parameter that differentiates between various occupations: the proneness to reality flukes – unexpected misestimations of true task difficulty. Many more examples of combining causal processes with agent-based modelling can be found in the companion modelling approach (ComMod) influenced by the French teams of Barreteau, Le Page, Bousquet, and colleagues (Barreteau et al., 2014).

Conclusions

We argue that agent-based modelling is an integrative platform for various types of causal relationships, as they are depicted by a given researcher. This extraordinary capacity to work with other methods is crucial when tackling the complex problems of the modern world. As representatives of generative social science and builders of agent-based models that explain macro-level social phenomena, we strongly believe that any feasible explanation of such phenomena cannot overstate the importance of the micro-level interactions through which the macro-level emerges. Explanations that actively ignore the interactions among individuals embedded in social networks or that deny these interactions matter, are in our view incomplete. However, in constructing our mechanistic, generative explanations we admit that not all causal relationships are multilevel, and even if a phenomenon can be seen as generated by interacting entities on a lower level, it is not necessary to always represent it as such. ABMs are capable of mirroring the crucial, multi-level component of emergent phenomena, while not limiting the scope of the offered explanation and recognizing the important role of single-level causal forces. Implementing various types of causal relationships to *complement* the generative causation offers insight about *how* a multi-level phenomenon happens and allows for building more complete causal explanations.

The causal relationships executed in an ABM can originally be expressed in various symbolic systems: in both human and computer languages, including various forms of scientific notation, singular linear and non-linear equations, systems of equations (e.g. present in multi-level or

structural equation modelling) and natural languages. Graphical expressions e.g. in the form of UML diagrams (Siebers & Davidson, 2015) or even theory of change diagrams (Wilkinson et al., 2021), can also be successfully used as ABM input. Challenges to translating the *original* language into ABM code may relate to assuring the modeller has sufficient insight into the original causal statement, especially if those statements are not expressed in scientific notation that can be straightforwardly implemented as model code. Therefore, in complex models it is crucial to cooperate with subject matter experts and take adequate time to guarantee not only the understanding of causal relationships going into the model, but also the consequences of the way those relationships were translated into the target ABM language. Besides that, any limitation for AMBs to include all relevant causal relationships and/or precisely quantify them is empirical and stems from challenges inherent to data collection and analytical techniques, rather than from the nature of an ABM *per se* and would equally apply to other modelling (and non-modelling) analyses.

Last, we tackle the micro-chauvinistic myth. Agent-based modelling is not committed *exclusively* to emergence, even though this type of causal mechanism is its distinctive characteristic emphasized by the Santa Fe Institute pioneers and vocal proponents who wanted to establish its methodological identity and differentiate it from other computational methods. As a stand-alone methodology, it is fully capable of representing bi-directional, multi-level causal relationships and single-level simple and complex causal inferences. Therefore, ABM immanently embraces various approaches to causality. Even though the macro-level outcome is indeed generative, experimenting in simulations follows a probabilistic, counterfactual approach (controlling for confounders by keeping multiple factors constant) and sensitivity analysis may follow a multiple conjunctural approach, looking for combinations of factors that impact the outcome of the model.

Notes

1. We thank Reviewer 3 for pointing out that multiple approaches to social mechanisms exist, including ones that are not micro-chauvinistic, incl. Bunge (2004) or Mayntz (2003).
2. The reflections on multi-level phenomena in section 2 are focused on two-level models i.e. micro and macro. In reality, ABMs are capable of representing more complex systems with multiple, partially-interacting layers of structure (Gimona & Polhill, 2011). Therefore, the interpretation of this section should not be limited to two level-models only.
3. The view has recently been contested, see, (Lange, 2017; Reutlinger & Saatsi, 2018).
4. For a review and critique of mechanistic theories, see, Williamson (2011).
5. In Machamer et al.'s (2000) version, complex-systems theory is a generalization of process theory.
6. For a broader discussion on types of emergence, see, Bedau (2002).
7. We thank Reviewer 2 for a thoughtful comment on the matter that challenged our initial thinking. Note that we follow a broad, functional definition of a research design: *The function of a research design is to ensure that the evidence obtained enables us to answer the initial question as unambiguously as possible* (De Vaus, 2001, p. 9).
8. For an alternative typology, see, Szostak (2015).
9. For other philosophical theories of probabilistic causality see, also Reichenbach (1956), Good (1961-1962), and Suppes (1970).
10. For a discussion on the controversial topic of causal explanation in qualitative methods, see, Maxwell (2004).
11. Or, perhaps, all the causal assumptions they know how to handle? – see, Edmonds (2015) on the use of qualitative data in ABM.

Disclosure statement

Notes on contributors

Patrycja Antosz is a sociologist and a psychologist whose research has focused on three areas: (1) social simulation of complex systems, (2) work and organization studies, and (3) public policy analysis. Having used agent-based modelling for over a decade, she specializes in applying the methodology in a mixed-method designs alongside qualitative and quantitative techniques. Patrycja is an active member of the scientific community, currently serving a second term in the European Social Simulation Association's Management Committee, and a firm believer in using science to support evidence-based policy making.

Timo Szczepanska is a PhD candidate at the Norwegian Fisheries College, UiT The Arctic University of Norway. With a background in Futures Studies and Philosophy, he works at the intersection of humans and technology, to support sustainable social and ecological transformation. In research, he combines agent-based modelling and games to represent and study social-ecological systems. Timo is a member of the Management Committee of the European Social Simulation Association and an advocate for interdisciplinary research.

Loes Bouman is a social scientist working at the University of Groningen. Her research and teaching focusses on interdisciplinary studies of individual's decision-making processes for participation in collective action and how groups achieve collective success. This by using and developing quantitative methods such as experiments and agent-based simulation models. In particular, she is interested in studying the role of prosocial behavior in the spreading of social innovations and opinion dynamics.

J. Gareth Polhill (known as Gary Polhill) did a degree in Artificial Intelligence and a PhD in Neural Networks before spending 18 months in industry as a programmer. Since 1997 he has been working at The James Hutton Institute on agent-based modelling of human-natural systems, and has worked on various international and interdisciplinary projects using agent-based modelling to study agricultural systems, lifestyles, and transitions to more sustainable ways of living. He is interested in complex systems, social dilemmas, spatially-explicit modelling, and best practice in agent-based modelling, programming and design.

Wander Jager is a complex systems scientist and agent-based modeller working at the University of Groningen. His research and teaching focuses on how interactions between people give rise to the emergence of collective behavior. In particular, he is interested in the spreading of new technologies, opinion dynamics & polarization and the societal transition towards sustainability. Wander also likes to collaborate with people from different scientific disciplines on interdisciplinary challenges, and is very curious of all scientific developments.

References

Antosz, P., Rembiasz, T., & Verhagen, H. (2020). Employee shirking and overworking: Modelling the unintended consequences of work organisation. *Ergonomics*, *63*(8), 997–1009. https://doi.org/10.1080/00140139.2020.1744710

Antosz, P., & Verhagen, H. (2020). Qualitative Data in the Service of Model Building - the Case of Structural Shirking. In H. Verhagen, M. Borit, G. Bravo, & N. Wijermans (Eds.), *Advances in Social Simulation: Looking in the Mirror* (pp. 33–45). Springer.

Archer, M. S. (1982). Morphogenesis versus structuration: On combining structure and action. *British Journal of Sociology*, *33*(4), 455–483. https://doi.org/10.2307/589357

Barnes, B. (2001). The micro/macro problem and the problem of structure and agency. In G. Ritzer & B. Smart (Eds.), *Handbook of sociological theory* (pp. 339–354). Sage.

Barreteau, O., Bousquet, F., Etienne, M., Souchère, V., & D'Aquino, P. (2014). Companion modelling: A method of adaptive and participatory research. In M. Etienne (Ed.), *Companion modelling: A participatory approach supporting sustainable development* (pp. 13–40). Springer.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of the Biological and Biomedical Sciences*, *36*(2), 421–441. https://doi.org/10.1016/j.shpsc.2005.03.010

Bedau, M. (2002). Downward Causation and the Autonomy of Weak Emergence. *Principia: An International Journal of Epistemology*, *6*(1), 5–50 doi:10.7551/mitpress/9780262026215.003.0010.

Bhattacharya, K., Takko, T., Monsivais, D., & Kaski, K. (2019). Group formation on a small-world: Experiment and modelling. *Journal of the Royal Society Interface*, *16*(156), 20180814. https://doi.org/10.1098/rsif.2018.0814

Bunge, M. (2004). How Does It Work?: The Search for Explanatory Mechanisms. *Philosophy of the Social Sciences*, *34* (2), 182–210. https://doi.org/10.1177/0048393103262550

Cedeno-Mieles, V., Hu, Z., Ren, Y., Deng, X., Adiga, A., Barrett, C., Contractor, N., Ekanayake, S., Epstein, J. M., Goode, B. J., Korkmaz, G., Kuhlman, C. J., Machi, D., Macy, M. W., Marathe, M. V., Ramakrishnan, N., Ravi, S. S., Saraf, P., & Self, N. (2020). Networked experiments and modeling for producing collective identity in a group of human subjects using an iterative abduction framework. *Social Network Analysis and Mining*, *10*(1), 1–43. https://doi.org/10.1007/s13278-019-0620-8

Clark, P. (1990). Explanation in Physics: Explanation in Physical Theory. *Royal Institute of Philosophy Supplement*, *27*, 155–175. https://doi.org/10.1017/S1358246100005099

Conte, R., Andrighetto, G., & Campennl, M. (2013). *Minding Norms: Mechanisms and Dynamics of Social Order in Agent Societies*. Oxford University Press.

Conte, R., & Castelfranchi, C. (1995). *Cognitive and Social Action*. UCL Press.

Conte, R., Edmonds, B., Moss, S., & Swayer, R. K. (2001). Sociology and Social Theory in Agent Based Social Simulation: A Symposium. *Computational and Mathematical Organization Theory*, *7*(3), 183–205. https://doi.org/10.1023/A:1012919018402

Conte, R., & Paolucci, M. (2014). On agent-based modeling and computational social science. *Frontiers in Psychology*, *5*, 668. https://doi.org/10.3389/fpsyg.2014.00668

de Vaus, D. A. (2001). *Research design in social research*. Sage Publications.

Edmonds, B. (2015). Using Qualitative Evidence to Inform the Specification of Agent-Based Models. *Journal of Artificial Societies and Social Simulation*, *18*(1), 18. https://doi.org/10.18564/jasss.2762

Elsenbroich, C. (2012). Explanation in Agent-Based Modelling. *Journal of Artificial Societies and Social Simulation*, *15* (3), 1. https://doi.org/10.18564/jasss.1958

Epstein, J. M. (2006). *Generative Social Science*. Princeton University Press.

Epstein, J. M., & Axtell, R. (1996). *Growing Artificial Societies: Social Science from the Bottom Up*. Brookings InstitutionPress; The MIT Press.

Filatova, T., Verburg, P. H., Parker, D. C., & Stannard, C. A. (2013). Spatial agent-based models for socio-ecological systems: Challenges and prospects. *Environmental Modelling & Software*, *45*, 1–7. https://doi.org/10.1016/j.envsoft.2013.03.017

Giddens, A. (1984). *The constitution of society: Outline of the theory of structuration*. University of California Press.

Gilbert, N. (2002). Varieties of Emergence. In D. Salllach (Ed.), *Agent 2002 Conference: Social agents: Ecology, exchange, and evolution* Chicago (pp. 41–56). University of Chicago and Argonne National Laboratory.

Gimona, A., & Polhill, J. G. (2011). Exploring robustness of biodiversity policy with a coupled metacommunity and agent-based model. *Journal of Land Use Science*, *6*(2–3), 175–193. https://doi.org/10.1080/1747423X.2011.558601

Good, I. J. A Causal Calculus I-II. (1961-1962). *British Journal for the Philosophy of Science*, *11*(44), 305–318. *12*, 43-51; *13*, 88. https://doi.org/10.1093/bjps/XI.44.305

Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S. K., Huse, G., Huth, A., Jepsen, J. U., Jørgensen, C., Mooij, W. M., Müller, B., Pe'er, G., Piou, C., Railsback, S. F., Robbins, A. M., . . . DeAngelis, D. L. (2006). A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, *198*(1–2), 115–126. https://doi.org/10.1016/j.ecolmodel.2006.04.023

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Hedström, P. (2005). *Dissecting the social: On the principles of analytical sociology*. Cambridge University Press.

Hedström, P., & Manzo, G. (2015). Recent trends in agent-based computational research: A brief introduction. *Sociological Methods & Research*, *44*(2), 179–185. https://doi.org/10.1177/0049124115581211

Hedström, P., & Ylikoski, P. (2010). Causal Mechanisms in the Social Sciences. *Annual Review of Sociology*, *36*(1), 49–67. https://doi.org/10.1146/annurev.soc.012809.102632

Hempel, C. G. (1965). *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*. Free Press.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, *15*(2), 135–175. https://doi.org/10.1086/286983

Holland, P. W. (1986). Statistics and causal inference (in theory and methods). *Journal of the American Statistical Association*, *81*(396), 945–960. https://doi.org/10.1080/01621459.1986.10478354

Hume, D. (1739 [1978]). *A Treatise of Human Nature*. Clarendon Press.

Lange, M. (2017). *Because Without Cause: Non-causal Explanations in Science and Mathematics*. Oxford University Press.

Lewis, D. K. (1973). *Counterfactuals*. Harvard University Press.

Lewis, D. K. (1986). *Philosophical Papers Volume II*. Oxford University Press.

Macal, C. M., & North, M. J. (2010). Tutorial on agent-based modelling and simulation. *Journal of Simulation*, *4*(3), 151–162. https://doi.org/10.1057/jos.2010.3

Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science*, *67*(1), 1–25. https://doi.org/10.1086/392759

Mackie, J. L. (1974). *The Cement of the Universe: A Study of Causation*. Clarendon Press.

Macy, M., & Flache, A. (2009). Social dynamics from the bottom up: Agent-based models of social interaction. In P. Bearman & P. Hedström (Eds.), *The Oxford Handbook of Analytical Sociology* (pp. 245–268). Oxford University Press.

Marilleau, N., Lang, C., & Giraudoux, P. (2018). Coupling agent-based with equation-based models to study spatially explicit megapopulation dynamics. *Ecological Modelling*, *384*, 34–42. https://doi.org/10.1016/j.ecolmodel.2018.06.011

Maxwell, J. A. (2004). Using Qualitative Methods for Causal Explanation. *Field Methods*, *16*(3), 243–264. https://doi.org/10.1177/1525822X04266831

Mayntz, R. (2003). *Mechanisms in the analysis of macro-social phenomena. MPIfG Working Paper*, 03/3, Max Planck Institute for the Study of Societies, Cologne, http://www.mpifg.de/pu/workpap/wp03-3/wp03-3.html

Mill, J. S. (1843 [1967]). *A System of Logic: Ratiocinative and Inductive*. University of Toronto Press.

Mohr, L. B. (1982). *Explaining organizational behavior: The limits and possibilities of theory and research*. Proquest Info & Learning.

O'Sullivan, D., & Haklay, M. (2000). Agent-Based Models and Individualism: Is the World Agent-Based? *Environment and Planning A: Economy and Space*, *32*(8), 1409–1425. https://doi.org/10.1068/a32140

Pawson, R., & Tilley, N. (1996). What's Crucial in Evaluation Research: A Reply to Bennett. *The British Journal of Criminology*, *36*(4), 574–578. https://doi.org/10.1093/oxfordjournals.bjc.a014112

Pérez-González, S. (2020). Mechanistic explanations and components of social mechanisms. *European Journal for Philosophy of Science*, *10*(35). https://doi.org/10.1007/s13194-020-00300-1

Polhill, J. G., Hare, M., Bauermann, T., Anzola, D., Palmer, E., Salt, D., & Antosz, P. (2021). Using Agent-Based Models for Prediction in Complex and Wicked Systems. *Journal of Artificial Societies and Social Simulation*, *24*(3), 3. https://doi.org/10.18564/jasss.4597

Pozzoni, G., & Kaidesoja, T. (2021). Context in Mechanism-Based Explanation. *Philosophy of the Social Sciences*, *51* (6), 523–554. https://doi.org/10.1177/0048393121991657

Ragin, C. C. (1987). *The Comparative Method: Moving Beyond Qualitative and Quantitative Strategies*. University of California Press.

Reichenbach, H. (1956). *The Direction of Time*. University of California Press.

Reutlinger, A., & Saatsi, J. (2018). *Explanation Beyond Causation; Philosophical Perspectives on Non-Causal Explanations*. Oxford University Press.

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, *21*(4), 25–34. https://doi.org/10.1145/37402.37406

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, *66*(5), 688–701. https://doi.org/10.1037/h0037350

Russo, F. (2009). *Causality and Causal Modelling in the Social Sciences*. Springer.

Sakoda, J. M. (1971). The Checkerboard Model of Social Interaction. *Journal of Mathematical Sociology*, *1*(1), 119–132. https://doi.org/10.1080/0022250X.1971.9989791

Salmon, W. C. (1971). Statistical Explanation. In W. C. Salmon (Ed.), *Statistical Explanation and Statistical Relevance* (pp. 29–87). University of Pittsburgh Press.

Salmon, W. C. (1984). *Scientific Explanation and the Causal Structure of the World*. Princeton University Press.

Salmon, W. C. (1997). *Causality and Explanation*. Oxford University Press.

Salmon, M. H., & Salmon, W. C. (1979). Alternative Models of Scientific Explanation. *American Anthropologist*, *81* (1), 61–74. https://doi.org/10.1525/aa.1979.81.1.02a00050

Sawyer, R. K. (2000). Simulating Emergence and Downward Causation in Small Groups. In S. Moss & P. Davidsson (Eds.), *Multi-Agent-Based Simulation* (pp. 49–67). Springer.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology*, *1*(2), 143–186. https://doi.org/10.1080/0022250X.1971.9989794

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton, Mifflin and Company.

Siebers, P., & Davidson, N. (2015). Engineering Agent-Based Social Simulations: An Introduction. *Journal of Artificial Societies and Social Simulation*, *18*(3), 13. https://doi.org/10.18564/jasss.2835

Suppes, P. (1970). *A Probabilistic Theory of Causality*. North-Holland.

Szczepanska, T., Antosz, P., Bernd, J., Borit, M., Chattoe-Brown, E., Mehryar, S., Meyer, R., Onggo, S., & Verhagen, H. (2022). GAM on! Six pathways to explore social complexity with games and Agent-Based Models. *International Journal of Social Research Methodology*.

Szostak, R. (2015). Interdisciplinary and Transdisciplinary Multimethod and Mixed Methods Research. In S. Hesse-Bieber & R. B. Johnson (Eds.), *Oxford handbook of multimethod and mixed methods research inquiry* (pp. 128–143). Oxford University Press.

Turner, J. (2001). *Handbook of sociological theory*. Springer.

Wilkinson, H., Hills, D., Penn, A., & Barbrook-Johnson, P. (2021). Building a system-based Theory of Change using Participatory Systems Mapping. *Evaluation*, *27*(1), 80–101. https://doi.org/10.1177/1356389020980493

Williamson, J. (2011). Mechanistic Theories of Causality Part I. *Philosophy Compass*, *6*(6), 421–432. https://doi.org/10.1111/j.1747-9991.2011.00400.x