Environmental Modelling and Software 168 (2023) 105802

Contents lists available at ScienceDirect

Environmental Modelling and Software

journal homepage: www.elsevier.com/locate/envsoft

What do you want theory *for*? - A pragmatic analysis of the roles of “theory” in agent-based modelling

Patrycja Antosz [a], Dan Birks [b], Bruce Edmonds [c,*], Alison Heppenstall [d], Ruth Meyer [e], J. Gareth Polhill [f], David O’Sullivan [g], Nanda Wijermans [h,i]

[a] NORCE Norwegian Research Centre, Kristiansand, Norway
[b] School of Law, University of Leeds, UK
[c] Centre for Policy Modelling, Manchester Metropolitan University, Aytoun Campus, Oxford Road, M15 6BH, UK
[d] School of Social and Political Sciences, MRC/CSO Social and Public Health Sciences Unit, University of Glasgow, UK
[e] Fraunhofer-Zentrum für die Sicherheit Sozio-technischer Systeme (SIRIOS), Kaiserin-Augusta-Allee 31, 10589, Berlin, Germany
[f] Information and Computational Sciences Department, The James Hutton Institute, Aberdeen, UK
[g] School of Geography, Environment and Earth Sciences, Victoria University of Wellington, New Zealand
[h] Stockholm Resilience Centre, Stockholm University, SE-106 91, Stockholm, Sweden
[i] Institute for Future Studies, Box 591, SE-101 31, Stockholm, Sweden

A B S T R A C T

There has been some discussion about agent-based modelling (ABM) and theory, particularly how ABM might facilitate theory building. However, there is confusion about the different ways they could relate and some scepticism as to whether theory is needed if one has an ABM. This paper distinguishes some of the different ways that the term “theory” is used in ABM papers in three important ABM journals: *Environmental Modelling & Software*, *Computers, Environment and Urban Systems* and the *Journal of Artificial Societies and Social Simulation*. Apart from the simple-minded identification of theory with mathematics, we distinguish nine different ways that theory and ABM relate. This analysis is situated with respect to some of the expectations and philosophical background behind the idea of “theory”. The paper concludes with some ways in which theory and ABM could work better together, some possible ways forward and suggests that a more cautious approach to generalisation might be more appropriate.

1. Introduction – why talk about theory?

There is considerable discussion about what the word “theory” entails, especially in the philosophy of science. This paper looks at the role of what is called “theory” in some of the agent-based modelling (ABM) literature, in order to extract some practical lessons.

However, before we talk about how “theory” is used first of all it is necessary to look at why we should bother talking about theory at all, since many ABM researchers do not mention theory but rather focus on models only. Of course, these researchers theorise, in the widest sense – interpreting their models in more general terms – but most do not label any of this as “developing a theory”. Often, they do not explicitly reference entities called “theories”, maybe because their models perform many of the roles that theory would have played. Indeed, when asked informally, many (including some of the authors here) might say that their theory *is* their model – identifying theory and their models – but many others do not claim their models encapsulate theories and are content with that.

However, other researchers do talk about theory and value it. Calling something a “theory” undoubtedly implies it has *status* – the label indicates that what it labels is worth considering as an entity. Furthermore, having theories is one of the hallmarks of a science, part of the paraphernalia one associates with a science (along with data, mathematics, conferences, journal papers etc.). Of course, these are merely indicators of a science; they do not necessarily make it scientific. Phrenology (the study of the shape of the skull as an indication of mental traits and character) had all the indicators of a science, including a pernicious theory, without any of its validity.

Lorscheid et al. (2019) call for theory development through agent-based modelling. This points out that ABM produces lots of specific models aimed at specific case studies but has not achieved more general explanations/patterns/theories from these – the so-called YAAWN syndrome (O’Sullivan et al., 2016), where everybody produces “yet another model” with little connection to other models and no apparent progress towards a more general understanding of the target systems. As a result, Lorscheid et al. (2019) suggest methodological

* Corresponding author.
E-mail address: bruce@edmonds.name (B. Edmonds).

https://doi.org/10.1016/j.envsoft.2023.105802
Received 30 April 2023; Received in revised form 5 July 2023; Accepted 17 August 2023
Available online 21 August 2023

development as part of a programme aimed at getting more theory out of our modelling. Previously, others have also seen ABM as a tool for developing theory. Carley (1995) foresaw a field using simulation tools which "… *has the potential to move theories of organizations beyond empirical description to generative formalizations*". Anderson (1999) also looked to ABM tools to develop organisation theory. Macy and Willer (2002) explained ABM as a strategy for building up theory. Hedström and Ylikoski (2010) looked to ABM to help develop mid-range theories in the form of causal mechanisms.

Thus, at least among some researchers, the idea of "theory" is important, and they hope that ABM will help develop theory in some way. Now, over 25 years since (Carley 1995), we look back on the issue to analyse how it has worked out, taking a pragmatic approach. We do not want to take a view about the issues but rather ground our discussion on a survey of how the term is used and what its role seems to be. Thus, this is not so much about the theory of "theory", but about how the term is used in practice.

In this paper, we first review some of the prior expectations of a "theory", briefly looking at the history of the relationship between theories and models in the philosophy of science, and the challenges for ABM researchers in dealing with theory. Subsequently, we move onto the main part of this article, reviewing papers in the fields of social, environmental modelling and the spatial sciences where ABM is used, in order to establish how the term "theory" is used. We summarise these uses and speculate about some possible ways forward before concluding.

2. Expectations of theories

If people claim some proposition as a theory or label their idea as a theory, this indicates something about the proposition to others. Here we briefly consider a little of the philosophy concerning theories in order to prepare the way for our main analyses. This is important as the history of what a scientific theory is sets the expectations for theory, regardless of whether these are met in practice.

In formal logic, a theory is the set of all statements that can be deduced from a set of axioms or assumptions, whilst a model is a structure that is consistent with these statements – it is a particular example of something that satisfies the theory (Tarski 1936). There is some confusion here, because there is another prior use of the word model, as in 'model aircraft' – something which is like, or acts like, what is modelled (Wartofsky, 1979). In the latter case there may not be any explicit theory, but, rather, the theory is implicit in terms of what characteristics and structures that were included in the making of the model.

In science, theory is often conceived of as being a general explanation of observed phenomena that has been empirically established, such as the theory of gravitation (Hempel 1942, 1965). The model here is a more specific realisation of a theory from which one is able to make testable inferences, but the model is not *generally* true because it is limited in scope, i.e. to a particular case or context (Braithwaite, 1962). In order to obtain a specific practical prediction or to be able to compare the theory to data one has to add in more facts from the case in hand to make a model, and it is the model that predicts. Although the idea that theories should be axiomatized as a set of logical statements (plus a mapping to observations) was attractive to some philosophers – the so-called "*Received view*" of theories originating with the Logical Positivists and cemented in the 1960 conference (Nagel et al., 1962), this is rarely done and, indeed, it is not clear how this would even be accomplished in most cases (Hempel 1974).

Whilst the debate about what a scientific theory *is* rumbles on in the philosophy of science, it is clear that how models and theories relate to each other and how science *actually* works with this relationship is important (Suppe, 2000). If we take a 'family resemblance' (Wittgenstein 1953) approach to "theory", then we might associate theory with many of the following characteristics: (a) it should have some level of generality – not necessarily global, but should be applicable to a set of phenomena and not only to one case (as in 'mid-range' theories), (b) it should include some kind of causal or mechanistic connection between its components in terms of which explanations are formulated, (c) it should be comprehensible or useable for humans, (d) it is usually at a relatively high-level description, (e) it is empirically reliable or replicable, (f) it can be used to infer things about the phenomena, (g) it newly posits and labels an entity (e.g. social capital), (h) it is formally expressed (e.g., using mathematics). However, not all of these hold for all examples of "theories" – for example, Darwin's "Theory of Natural Selection" was not able to provide testable conclusions, until the new-Darwinian synthesis with genetics in the early 20th Century.

This family resemblance leaves ABM researchers with a problem. People talk about theory for different reasons and on top of that, it means different things to different researchers. Theory often poses severe challenges for ABM researchers: (1) it can be very vague and ill defined, (2) it may not constrain possible interpretations very much, having a low explanatory content, (3) it might be founded on a great many assumptions, (4) it may assume a certain context of application that may not be explicitly described, (5) it may require many further assumptions to make it testable, (6) it may be more of a way of thinking about stuff, an analogy, rather than have any causal or mechanism-like content (e.g., the idea of "resilience"). One can understand why some prefer to avoid the term "theory" altogether and just deal with models.

To avoid these definitional problems, we will not take an essentialist view as to what a theory *is*, nor a normative approach of what it *should* be, but rather restrict ourselves to how the word "theory" has been used in ABM papers, and then try to extract some practical conclusions from this. This is a deliberately inductive method which may frustrate some readers who are used to a more opinion-led approach. In other words, this paper is not about the theory of theory but, rather, its practice.

3. How the word "theory" is used in some ABM fields

Here we give examples from ABM papers that talk about theory *explicitly*. Not so many papers talking *about* ABM or the future prospects for ABM and theory but those showing examples of actually doing it. We look at three fields: social simulation, environmental modelling and the spatial sciences.

Using the expertise present in the team of authors, we identified a renowned journal with a history of publishing agent-based models in three fields: environmental modelling (*Environmental Modelling & Software*), spatial sciences (*Computers, Environment and Urban Systems*) and social sciences (*Journal of Artificial Societies and Social Simulation*). On the January 25, 2022 we performed a systematic search for articles that include an agent-based model and maintain a focus on theories (search terms "agent-based" and "theory*" in title, keywords or abstract). We covered all papers published by these journals during the time they have existed. Out of the articles that matched the search criteria, we excluded works that did not present an ABM. Furthermore, although a reading of some papers might *imply* theoretical uses, we focussed down on those that explicitly use the terms: "theory" "theories" or "theoretical" The results of the search are summarised in Table 1, below.

In each of the sub-sections below (one for each field considered), we first briefly review the discussion in that field about theory in ABM, before turning to a consideration of the articles in the corresponding journal that (a) present a specific agent-based model and (b) mention the word "theory" (or theories).

3.1. 'Theory' in environmental modelling papers

In the following, we constrain environmental modelling to the application of ABM to study social-ecological systems (SES). Due to their

[1] We used a script that scanned JASSS articles automatically in 2021. This found 127 such papers.

Table 1
A summary of the three chosen review fields the journals focussed upon and the number of articles reviewed.

Field: Journal Title	Environmental modelling: Environmental Modelling & Software	Spatial sciences: Computers, Environment and Urban Systems	Social sciences: Journal of Artificial Societies and Social Simulation
Number of articles matching the search terms	13	23	117
Number of articles qualifying for the review	5	17	>36[1]
Number of articles reviewed	5	17	36

innate ability to represent heterogeneity in both actors and their interactions, ABM have been widely used to investigate human-environment interactions with the aim to advance theory and inform policy making in natural resource management (Schulze et al., 2017). Formal modelling in the form of ABM has helped to develop our knowledge of the dynamics of natural resources, their response to management interventions and environmental change, as well as their vulnerabilities and regenerative capacities (Schlüter et al., 2017).

One of the main challenges has been – and still is – the adequate representation of human decision-making. There is "no culture of rigorous theory development, which would require that alternative representations be implemented and tested for their ability to reproduce multiple patterns observed in real social systems" (An et al., 2021). Most models of SES base their agent decision-making on ad hoc assumptions (Crooks et al., 2008, Smajgl and Barreteau, 2014) or oversimplify human behaviour by adopting the standard view of economic theory that all actors are selfish and rational. A recent survey (Groeneveld et al., 2017) of 134 agent-based land use models found that the majority (62%) of these models did not explicitly ground their human decision-making processes on theory, whereas the most frequently applied theory was Expected Utility Theory (35%), a variant of rational choice under uncertainty, followed by Satisficing (9.7%). This "certainly reflects the dominance of economics compared to psychology in the field" (Groeneveld et al., 2017). What is interesting (or rather alarming) is that there has not been an increasing attempt to base the human decision-making component on theory.

3.1.1. Analysis of ABMs in Environmental Modelling & Software

To investigate the use of theory in ABMs of socio-ecological systems we undertook a survey of *Environmental Modelling and Software*, one of the most relevant journals in the area. Of the 13 papers identified by the search described above, 6 had to be excluded for not containing an ABM. Another one proposes an extension to the established ODD protocol to include human decision-making in the model description with a section on the theoretical background (Müller et al., 2013), while one paper gives a systematic review of the theoretical foundations for human decision-making in agent-based land use models (Groeneveld et al., 2017). The remaining five papers cover applications ranging from agriculture and land use change (Pacilly et al., 2019; Coelho and Ralha 2022; Barnaud et al., 2013) to natural resource management (Touza et al., 2013) and climate change mitigation (Niamir et al., 2020).

The majority of these papers use one or more theories from outside their application domain to help specify a component of the ABM, usually the decision-making of the agents. Evolutionary game theory is the basis for land manager agents to decide how many deer to cull per season in the agent-based model of natural resource management by Touza et al. (2013). Based on the received payoff and the comparative success of their neighbours, agents can adapt their strategy (proportion of deer to kill) for the next time step. An interesting aspect of this model is that the natural resource (deer) is mobile, so payoffs may change due to the number of deer present on the land.

Coelho and Ralha (2022) apply evolutionary game theory to model agent interaction and conflict resolution, arguing that the emergence of social phenomena like certain behaviours in human society can be interpreted as evolutionary games. In their model agents competing for land parcels repeatedly play 1-1 games until the winner is determined and allowed to decide the land use for the parcel in question. Losers are able to adapt their strategies by learning from the outcomes of the games and either take on the strategy of their opponent (reactive mode) or the strategy with the currently highest payoff (registry-based mode).

The other two papers use behavioural theories from psychology to inform the decision-making of their agents. Pacilly et al. (2019) rely on the Consumat framework (Jager and Janssen 2012) to model farmers' decision-making regarding disease control in potato plants. Originally devised to model consumer behaviour (Jager et al., 2000), the Consumat approach incorporates aspects from a variety of psychological theories on human needs, motivational processes, social comparison, social learning and reasoned action into a comprehensive conceptual framework, answering the call for a meta-theory of human behaviour (Vallacher and Nowak 1994). Key drivers for the selection of behaviour are need satisfaction and uncertainty, resulting in agents applying one of four strategies to determine which behavioural options to adopt: repetition, imitation, inquiring and optimising.

To overcome the limitations of macroeconomic models in representing heterogeneous agents and behaviour change, Niamir et al. (2020) combine an empirically calibrated computable general equilibrium model (CGE) (Ivanova et al., 2019) with an agent-based model of household energy choices (Niamir et al., 2020). The use of theory here is two-fold: Firstly, the agent-based model draws on established psychological theories (theory of planned behaviour (Ajzen 1985), norm activation theory (Schwartz 1977) and value-belief-norm theory (Stern et al., 1999)) to model the decision-making process of its household agents. Secondly, by aggregating individual choices and feeding them into the CGE model, the ABM is used to refute the assumptions of classical economic theory about human behaviour: that rational, homogeneous actors are solely seeking to maximize their utility.

The "odd one out" is the paper by Barnaud et al. (2013), which describes a companion modelling process using role-playing games and agent-based models to facilitate discussions between stakeholders involved in the establishment of a new national park in Thailand. The use of theory here is not directly related to the ABM but to the negotiation process evolving around different model scenarios. When focussing on the proposed boundary of the national park discussions tended to be distributive, i.e. participants 'divided up the cake' (access to land/resources) with the most vulnerable stakeholders (poor farmers) set to lose out, while an – admittedly unrealistic – scenario without fixed boundaries let the stakeholders see options for joint uses of the land in question, thus resulting in an integrative mode of negotiation where participants try to reframe the problem to 'enlarge the cake'.

3.2. 'Theory' in spatial science papers

Typically, within the spatial sciences, theory has been developed based on sparse data and observations around phenomena such as migration, diffusion of ideas/products and gentrification ('classic' theories are typified by those put forward by Hägerstrand, 1952; Christaller, 1933). Many of these theories were derived at a time when fine grained data was largely unavailable and computational resources were sparse. These spatial theories were constructed from a largely deductive perspective based on common-sense assumptions such as Tobler's law – that "everything is related to everything else, but near things are more related than distant things." (Tobler 1970, page 236).

The famous Schelling-Sakoda model of segregation (Schelling 1971; Sakoda 1971) is able to formalise the processes described in these. It shows that it is possible that "… marked segregation can arise from rather mild individual preferences for living amongst one's own kind" (Crooks 2010) – and is thus a counter example to the assumption that segregation must result from a strong preference for what is perceived as one's own kind. But this model is highly abstract and is being used in this example to develop ideas about residential segregation, rather than to develop theory per se. Conclusions of the paper state that the model developed represents individuals and their preferences at the micro level with "recognisable patterns emerging at the macro scale", which is something like a theory, however this isn't quite the same as developing or testing theory.

There are many ABMs that develop the Schelling-Sakoda approach (Sakoda 1971; Schelling 1971). Crooks (2010) investigates the effect of moving away from the rectilinear grids used by Schelling to arbitrary (vector) shapes, facilitating relating this kind of model to GIS data. He sees this as a way of testing the robustness of its underlying assumptions, and clearly sees this as a start of a process of moving towards something like theory, saying "The model provides the essential ingredients for cumulative scientific inquiry with a clearly specified model that facilitates replication and extension" (2010, page 674). However, changing the model in this way also reveals a host of new issues concerning scale and the interpretation of residential clustering. Picascia (2017) adapts a Schelling kind of model but bases the micro-level changes in rent gap theory. He shows that this can result in some plausible macro-level behaviour and that the model can reproduce changes in price in both Manchester and London to a convincing degree. However, there are many other elements in his model so this can only be seen as a test of the combination of many theories and assumptions. If a "theory" is a series of models with strong family resemblance (Giere 2010) then the constellation of Schelling-Sakoda like models might qualify.

Since that time, ABM has been successfully used for decades to replicate and simulate a range of spatial models from very abstract (e.g., Schelling 1971; Magliocca et al., 2011) to more empirical models with a a focus on prediction (e.g., Bonabeau 2002; Hoertel et al., 2020; Rosés et al., 2021) and everything in-between depending on the models' purpose (Wu 2002; Epstein 2008). Perhaps the most abundant of spatial ABMs are for the representation of micro-scale processes, such as individual movement around neighbourhoods or city centres. Here, the ability of ABMs both to handle dynamic complexity and a range of data types can be readily exploited (Helbing and Balietti 2011).

Several modelling approaches show that introducing heterogeneity to agent preferences makes a significand difference to the outcomes, including Parker and Meretsky (2004) and Sasaki and Box (2003) which demonstrated how spatial patterns can result from economically rational agent behaviours. These seem to aim towards something like a theory (or testing a theory) but remain very abstract. They do not seem to have the power to refute a theory, for example, due to the necessary presence of other assumptions.

However, despite the success of ABMs, there is criticism about the way in which theory is embedded/represented by these models. Crooks et al. (2008) commented that while most models contain theories, the "*theoretical implications of many ABMs remain implicit and hidden, often covered by a thick veil of* ad hoc *assumptions about structure and process*". This is a point that O'Sullivan et al. (2016) develop in their critique of ABMs commenting that "*… there has been a drift away from using ABMs to engage with theory … Instead an increased focus on applications has directed attention to more ad hoc efforts attempting to build realistic models of particular systems.*

The remainder of this section will examine examples from urban analytics - a particular, well-developed area of the spatial sciences - to look at how theory is embedded within these models and whether the role of theory is explicit within the model or whether the theory is dealt with in an ad hoc manner.

3.2.1. Analysis of ABMs in Computers, Environment and Urban Systems

In this subsection we review the use of "theory" in one of the most popular applied geography journals, "Computers, Environment and Urban Systems". The papers reviewed covered a range of applications from urban traffic dynamics, crime to land-use modelling. Of the 17 papers reviewed, 10 came under the label of (1) when a theory is imported from elsewhere to help specify/justify/label a component of a wider ABM with only 2 papers under (2) the focus of the whole ABM is an existing theory or an idea.

Applications that exemplify this mainly fell within the area of land use modelling (e.g. Koomen et al., 2015; Zhuge et al., 2016; Robinson et al., 2012). Here, utility theory or rent gap theory is used to validate the processes contained within the ABM and decisions taken in designing the ABM. These examples do not seek to test the theories, but instead use them to build an aspect of 'realism' into the application. In terms of individual behaviour, aspects of theory can also be found embedded in models of crime. Both Rosés et al. (2021) and Zhu and Wang (2021) exemplify this through use of Routine Activity Theory in their design and justification of agent behaviour to predict the occurrence of different crime events.

Whilst the majority of applications use theory in an ad hoc manner, there are a couple of examples where the focus goes a little deeper i.e. the whole ABM is an existing theory or idea. Filatova (2015) uses economic theory combined with rich spatial data and econometric analysis to build an ABM of land markets whilst Liu and O'Sullivan (2016) use several theories (rent gap theory, filtering theory and household life cycle theory) to build a detailed model of gentrification. Whilst the previous examples use theory to design and validate part of the model, these examples have the theory as the core of the model.

Huang et al. (2013) present a detailed overview of how ABM has been used in the development of urban residential models - for example, several published ABM studies (e.g. Torrens and Nara, 2007) use 'representative features' drawn from bid rent theory. Here, the embedding of theory within ABM seems to be more of a half-way house i.e. only elements of a theory appear in the form of variables, but the treatment of theory is more implicit than explicit.

The above examples suggest that the theory incorporated is only tested along with a host of other assumptions. On the other hand, new theory coming out of ABM spatial simulation is more of a prospect than an actuality, at least as yet (echoing Crooks et al. (2008) comment above that theory in models is hidden under ad hoc assumptions about model structure and process). If theory can be a family of closely related models (Giere 2010) then indeed 'theory' can result in this sense.

3.3. 'Theory' in social simulation papers

Many in the field of social simulation do not bother with theory, except where the label is used for ideas imported from other fields. Others equate theory and their models, saying, if pushed, that their model *is* their theory[2]. Where theory is discussed more generally it is often a proxy for discussions about modelling methodology and use, e.g. as a route to importing relevant arguments from philosophy (e.g., Troitzsch 2009; Elsenbroich 2012). Some use their models to attack specific theories, e.g. the overly strong assumptions behind much economic theory (e.g., Moss 2008).

However, as in other fields, many researchers in the field feel a need for theory (e.g., Carley 1995; Anderson 1999; Macy and Willer 2002; Lorscheid et al., 2019). It is difficult to pin down this feeling, but maybe being stung by the criticism of its relative absence compared to the 'hard' sciences, they want to prove that they are a science or maybe the success in using abstract models illuminating ideas means that theory seems to be within their grasp.

[2] This is not said explicitly in their papers, but if one 'corners' them in a workshop this is what they say.

3.3.1. Analysis of ABMs in the Journal of Artificial Societies and Social Simulation

'Theory' was mentioned many times within the Journal of Artificial Societies and Social Simulation (JASSS) in ABM papers, but this is not surprising as JASSS is particularly focussed on ABM and has always included methodological as well as more applied papers. To ensure that this paper is not too imbalanced towards JASSS, we thus limited the review to 30 papers, composed of the 6 papers that mention 'theory' the most (Vu et al., 2020; Taghikhah et al., 2021; Dilaver 2015; Balzer et al., 2001; Spaiser and Sumpter 2016; Hassan et al., 2013) plus the 24 most cited papers that do (there was no overlap). We start with the papers that use the terms prolifically and then (more briefly) look at the others.

One of the papers mentioning the word 'theory' most is Vu et al. (2020), which presents a modelling framework within which ABMs can be developed that support a "middle range theory approach" aiming to discover possible mechanism-based explanations. In paragraph 1.13, they hint at their interpretation of what a theory is in a mereological statement justifying the requirement for a common language in the architecture the article describes to express "the diverse range of *mechanisms* and *entities* that constitute a theory" (emphasis is ours); noting in the ensuing paragraph that "a theory sometimes only explains a limited aspect of an observed phenomenon". This is a framework to support the implementation of ABMs from existing theories. The plurality of potential implementations of a theory in computer code is acknowledged (see Muelder and Filatova 2018), which showed that different implementations of the same theory can result in significantly different outcomes. Even so, they argue that social simulations founded on theory are preferable in that they "are not ad hoc, but rooted in sociological discourse" (para. 5.1) - and thus avoids the problem of simulators just making stuff up on the basis of common sense or to fit the data. In both of these, theories come from elsewhere and (partially and incompletely) these provide external constraints for the design of the ABM to avoid this problem.

Founding social simulations on theory, and the 'contrast' between theory-driven and empirical models, is the subject of Taghikhah et al.'s (2021) study, which compares a version of a model in which the agents' behaviours are based on formalizations of social theories, with a version of the same model in which the behaviours are developed using statistics and machine learning algorithms from data. Whilst there are significant differences in the result of these two models, in this case both fit the shift to organic consumption as accurately as the other. In the same sentence (para. 1.6), they claim that "theory-driven models are powerful tools in representing general dynamics", while quoting Sun et al.'s (2016, p. 57) observation that "simple theoretical assumptions" are often used in models as a cover for the *lack* of empirical knowledge and data. Just as Vu et al. (2020) do, Taghikhah et al.'s (2021) theoretical model combines theories – Ajzen's (1985) theory of planned behaviour, alphabet theory (Zepeda and Deal 2009), goal framing theory (Lindenberg and Steg 2011), and cognitive dissonance (Festinger 1962). This emphasises Vu et al.'s (2020, para. 2.19) points about theories' individual insufficiency when faced with preparing the algorithms needed to develop a model in a particular context. It also raises the question of whether and how to integrate theories in a single model (which connects to Voinov and Shugart's (2013) cautionary article about model integration in Environmental Modelling and Software).

Almost all of the usages of theory in Dilaver (2015) refer to 'grounded theory', a systematic methodology to develop theory based on the analysis of (qualitative) data from the bottom up - inductive reasoning (Glaser and Strauss 1967). It is worth noting Dilaver's (2015, para. 1.5) commentary on Glaser and Strauss (1967) being positive about the use of theory, in contrast with Vu et al.'s (2020, para. 5.1) negative opinions on models not founded on theory cited above. Dilaver expounds a more inductive approach to model building, starting from qualitative data, noting the common concern that "Both social simulation models and grounded theory research often receive scepticism, if not severe criticism, about whether or not their methods would accommodate *anything goes*." (Dilaver 2015, para. 1.5 – emphasis the author's.). The themes and patterns resulting from a grounded theory can be seen as a cautious approach to abstraction, but one that results in patterns and themes rather than anything causal. These patterns are then used to inform the design of the ABM.

The debate about whether models should be founded in theory goes back to the earliest years of *JASSS*, and is given detailed general attention by Balzer et al. (2001), albeit in the context of a specific (and, as the authors give the impression of believing, rather tedious) argument at the time between game theorists and social simulation modellers about which is the better approach to understanding social phenomena. In section 3 (paras. 3.1–3.16), the article addresses the apparent criticism from the game theorists' side that simulations lack an 'underlying theory' by breaking down the constituents of a 'scientific theory' (para. 3.5) into: hypothesis, data, specification of intended system and community of practitioners. It then translates the hypothesis and data into the simulation domain (para. 3.10) as being the model's computer program and initialisation data. While acknowledging (para. 3.14) that a simulation model may have only a single practitioner (the developer), they are then explicit in making the claim that "every simulation study has the same status as a scientific theory" when it is first proposed (para. 3.16). Founding a model in theory assumes that there exists a theory that is adequate to the task. "Game theory" is not a single theory, but rather an approach that has produced a set of models with a strong family resemblance (Giere 2010). Game theory is thus more a framework for modelling interaction – inference can be gained from a model that adds sufficient detail but not from the framework itself. In a different domain, Spaiser and Sumpter (2016) give specific attention to Human Development Sequence theory (Ingelhart and Welzel, 2005). In their paper, they find that they need to augment the original theory about how democracies emerge to account for social inequality in access to resources (Spaiser and Sumpter 2016, para. 5.26).

Many of the usages of the term 'theory' in Vu et al. (2020) reference existing work that is typically written with the word 'theory' – e.g. 'norm theory' (Rimal and Real 2005; Cialdini et al., 1991), 'role theory' (Knibbe et al., 1987; Wilsnack and Cheloha 1987; Yamaguchi and Kandel 1985), 'rational-choice theory' (Becker and Murphy 1988), 'theory of planned behaviour' (Ajzen 1991). In such contexts the status of the concept referred to as 'theory' is unquestioned; rather, it is simply a token to refer to a body of knowledge already accepted by one community or another as having the status of being a theory. Not using the word 'theory' to refer to the body of knowledge would then seem strange, and the usage is merely a matter of habit. Less unquestioning usages in Vu et al. (2020) are categorical in nature, including 'middle-range theory', 'single-theory', 'multi-theory', 'mechanism-based theory'. To these, Dilaver (2015, para. 3.27) adds 'mainstream theory', Balzer et al. (2001) add 'scientific theory'. However, Vu et al. (2020) also refers to discipline-oriented categories – 'psychological theory', 'social theory', 'cognitive theory' – which are more akin to referential usages, even though the constituents of a theory and criteria for recognition of knowledge as such might well differ from one discipline to another.

Spaiser and Sumpter (2016) and Vu et al. (2020) make reference to macro-micro-macro dynamics in their articles, with references to Hedström and Swedberg's (1996, 1998) metamodel, and Coleman's (1986) 'boat'. A key advantage of agent-based models over other approaches is the explicit representation of micro-micro emergent behaviour (micro-macro) and that, at least in principle (if not always in practice) agents could observe and reason about the macro level to inform their action, or have their behaviour constrained by macro outcomes (macro-micro) (Hassan et al., 2013). However, many of the specific theories referenced pertain to micro-micro interactions and do not include macro-micro processes. The arguments given by Balzer et al. (2001) already put pressure on the idea that models *must* be founded on theory to be scientifically sound. However, one point Balzer et al. (2001) does not raise in objection to this principle, which is still alive and well nearly twenty years later in Vu et al.'s (2020) article, is the path-dependency this creates: the set of theories available at any one time is a function of the work done by theoreticians. Historical (and to a significant but proportionally lesser extent ongoing) disciplinarity of such studies means that there is variation in the size of the menu of theories for each part of the 'boat' in any specific modelling case study. In social-ecological systems modelling the matter is brought into even sharper relief there, because each of the macro-micro-macro links also has social-social, social-ecological, ecological-ecological and ecological-social dimensions.

Moving to the 24 most cited that mention 'Theory', many of them used the term when a theory is imported into a simulation having already been called a theory (either individually or under a collective label), most frequently the term 'game theory' (Galan and Izquierdo 2005; Izquierdo et al., 2008; Hartshorn et al., 2013), with Rauhut and Junker (2009) also mentioning boundedly rational alternatives, and (Power 2009) also mentioning 'fuzzy set theory'. Two of these (Izquierdo et al., 2008; Rauhut and Junker 2009) applied the theory in the sense of using the mathematical machinery of game theory to infer conclusions (as well as via ABMs), the rest simply used the structure of game theory in terms of how interaction was implemented. Other theories that were mentioned in the selected papers included 'activities theory' (Fonoberova et al., 2012), 'costly signalling theory' (Wildman and Sosis 2011), 'self-categorization theory' (Salzarulo 2006), 'optimal distinctiveness theory' (Smaldino et al., 2012), 'theories of mobilisation' (Armano et al., 2003), 'theory of planned behaviour' (Knoeri et al., 2011), the 'Act-R memory model' (Wijermans et al., 2013) elsewhere called a theory in the paper, various theories from environmental psychology (van der Kam et al., 2019) and theories of emotion and computation (Staller and Petta, 2001). In seven of these 'theory' (Galan and Izquierdo 2005; Izquierdo et al., 2008; Fonoberova et al., 2012; Wildman and Sosis 2011; Hartshorn et al., 2013; Salzarulo 2006; Staller and Petta, 2001) only played a part in terms of labelling the justification for a core modelling element and the term had no other usage or significance in them.

The second most common usage among the 24 papers was to designate a more general framework, within which the implemented ABM was specified. This was less precise and tended to indicate more of a way of approaching the subject matter. Thus (Bergman et al., 2008) took the general framework of 'Socio-Technical Transitions' (Lustick 2000), within 'constructivist theory' (Holtz 2014), was informed in a general way by 'practice theories' (Rand et al., 2015), claimed to be 'grounded in social theory', (Knoeri et al., 2011) used the general ideas of Giddens' 'structuration theory'. This is a matter of degree rather than absolute kind – these took a vaguer and more general inspiration from a collection of ideas (grouped as a 'theory'), whilst in the previous case the 'theories' justified the specification of a particular mechanism or process. However, some of the latter were not described precisely in its original form and were, at least, somewhat vague.

More ambitiously in a few papers the whole of the ABM was an implementation of a theory (as opposed to a specific component or a general approach). Gilbert et al. (2001) seek to make a theory of innovation networks – a whole modelling framework instantiating and specifying this theory, which will then be applied to different cases. The whole of the ABM in Li and Xiao (2017) is an instance of 'social judgement theory'. (Muelder and Filatova, 2018) compare different versions of the 'theory of planned behaviour'. A whole approach to using parameter exploration to 'test' a theory (implemented as a sub-model) is described in (Thiele et al., 2014) with example ABMs, each implementing the theory to be tested.

Using an ABM to result in theory (other than simply implementing it) were rarer, though this was the *implicit* goal (e.g., of future work) implied in many of the papers already mentioned. The most common of these was to refine the theory by implementing it and revealing the additional assumptions or choices needed to get there. This included (Holtz 2014) for practice theories, the 'optimal distinctiveness theory' of social identity in (Smaldino et al., 2012) – pointing out that these will never be resolved at the purely individual level, the versions of the theory of planned behaviour (Muelder and Filatova, 2018), and to some extent (Rauhut and Junker 2009) in terms of the need for refined theories.

Two papers aimed to test theory using an ABM (Rand et al., 2015): on diffusion mechanisms in social media, and (Wijermans et al., 2013) that aimed to produce 'testable theory' in terms of their multi-level analysis of ABMs. However, it is clear in both of these that many other assumptions were also imported into these models to be tested together with the headline theory, which is OK as long as these other assumptions are not critical to the results.

One paper aimed to show that one input theory explained the observed data better than another given the other constant things in the models (van der Kam et al., 2019). One paper had elements of an inductive move towards new theory by observing the qualitatively different effects of theory elements upon the outcomes (Power 2009).

Finally, three papers associated theory with mathematics (Izquierdo et al., 2009): with the mathematics of Markov chain analysis, (Staller and Petta, 2001) mentioned deontic logic being some of the theory behind social norms and (Rauhut and Junker 2009) with the maths of game theory that they used.

4. Theory as mathematics

There is a tradition (coming from physics and hence economics) that simply calls any analytic mathematics, "theory" and any simulation, a "model". Where by "analytic" we mean it is possible to solve the mathematics to obtain a general "closed form" equation that specifies the value of output variables over the possible (maybe using appropriate approximations). Thus, an analytically solvable mathematical model can take the form of a general, theory-like formulation. However, the mathematics of most complex systems are not analytically solvable, which means that either (a) the maths only addresses a very restricted aspect of the system (e.g. when it is in equilibrium) or (b) they involve strong simplifying assumptions. Thus, with complex systems, the apparent generality of analytic solutions may be deceptive as the solutions obtained pertains to either: only the narrow aspect that the maths addresses, or only when the simplifying assumptions hold. ABMs can avoid these restrictions because it calculates solutions rather than using analytic proof but, on the other hand, each simulation run gives the output only for a specific set of inputs and random seed, so one can only sample the set of all outcomes.

Whilst equating theory with maths might be understandable from a historical perspective, it does not seem to be very useful terminology in today's context. It confuses the tool (the kind of formal system) with what it expresses – any maths could be made into a computational model and *vice versa* but this would not change the generality of its content. Maths *can* express general patterns but might also be used for describing very particular cases – there exists more theory-like and more model-like mathematics. Similarly, specific simulations can be summarised by more theory-like simulation models. We already have the words "simulation" and "mathematics" to distinguish these two so this maths = theory usage does not allow new distinctions and indeed confuses methodological considerations. Thus, although there are examples of this usage in ABM papers, we do not find it very helpful.

5. What theory goes "in" to an ABM

Although it is far from the case that all ABMs use anything called a theory in their formulation, they almost always are based on some ideas that are theory-like.

First, we look at when theory is used in construction of an ABM. That is in terms of informing its specification, before the results are known. This is irrespective of whether what is specified is about the macro-level (as in statistical or traditional economic modelling) or the micro-level (as in generative modelling). What we are distinguishing here by "in" and "out" is what is assumed compared to what is inferred.

We distinguish several cases, as follows.

(I1) *When a theory is imported from elsewhere to help specify/justify/label a component of a wider ABM.* Here the aim is to restrict choice – to constrain how the ABM is implemented – especially in terms of agent behaviour. This avoids the danger of entirely ad hoc models and has the happy side effect of justifying an element of an ABM within a published paper. However, the effectiveness of this approach depends upon how vague the theory is (and thus the range of implementations that are consistent with the theory) and the reliability of the theory being used (in other words, its empirical support).

(I2) *When the focus of the whole ABM is an existing theory or idea so the purpose of making an ABM from it is to test/refine this* – often bridging between micro and macro levels. In this case (as well as the one above), what results is usually some understanding of the results of a combination of the focus theory, but in combination with a whole host of other theories and assumptions – in this case, it does not convincingly test the focus theory but rather explores its potential outcomes with respect to this set of other assumptions. This has the potential to identify gaps as a result of formalisation (Sawyer, 2005), but, in many cases, the number of possible ways of filling these gaps is very large.

(I3) *The specific ABM is designed within a more abstract framework* (sometimes called a 'theoretical framework'). The impact of this approach depends upon how much the framework constrains what can be implemented within it. A strongly constraining framework makes this effectively similar to the *I1* or *I2* cases, whilst a weakly constraining framework becomes more of a programming language which is used in implementation – making different models easier to compare (which has the potential to facilitate generalisation to a theory, however we did not find any examples where this was achieved).

Clearly these categories can overlap. For example, one might take a theory from elsewhere, intending to use it to specify an ABM but find that one has to refine it in order to do this. The point is that each of these has different goals.

6. What theory comes "out" of an ABM

Is it not necessarily the case that we want any theory to result from a simulation – a major outcome of ABMs is not about theory at all but exploring the consequences of a particular set of entities, processes, structures and settings. However, there are several kinds of expectation when one does. The first four cases are as follows.

(O1) *That one might refute or support a theory of the (I2) kind above.* This does not give a straightforward refute/support example, but degenerates into the *O2* type discussing what kinds of assumptions might be necessary for the theory to hold.

(O2) *It might result in a refined theory* (or set of theories) of the *I2* kind demonstrating a concrete instantiation of the original (more vague) theory, revealing the additional assumptions that make this possible. This does seem to work as stated, but there are often alternative possible instantiations of the original theory.

(O3) It might show that one theory of the I1 kind explains observed outcome patterns whilst another (or a null hypothesis) does not. There are some examples of this, but the conclusions about which I1 theory is right are usually relative to a whole host of other assumptions that went into the ABM, so the conclusion is only as reliable as those assumptions.

(O4) *The ABM is used in an inductive fashion to suggest a completely new theory.* We have not come across many examples of this – it seems to be currently more of a hope than realised.

There were two further 'theoretical' uses of ABMs that were not explicitly mentioned in the papers reviewed but were implicit in many of them. These are as follows.

(O5) *When an ABM can show the consistency of a complete set of theories and assumptions with each other* (and hence enhances their plausibility). Maybe this is simply assumed by simulation modellers and thus not worthy of specific comment. Is a common and productive use, but only allows weak theoretical conclusions (in the sense that it is difficult to infer anything useful from them) depending upon how often that combination of theories and assumptions go together (and this is often not part of the main purpose of such integration exercises).

(O6) *When a family of related models results from a body of work, collectively constituting a 'theory' in some sense,* as Giere (2010) argues. Theories of this latter kind tend not to be trumpeted within a single paper but is maybe something that emerges within post-hoc reviews of the research landscape. This is perhaps the most convincingly achieved of the possible results of ABM modelling, but the establishing the relationship between the explored abstract families of models and empirical models or data has, up to now, not been extensively focussed upon.

7. Some possible ways forward

This brief survey has highlighted substantial holes in terms of effectively testing existing theories or generating new theories. If ABM models and theory are to work more productively together, as anticipated in some of the more visionary papers mentioned in the introduction, then we think substantial methodological advances are needed (hopefully supported by the development of new tools to support these). This will require a lot more rigour in how we use models and precision in terms of what one wants the theory to do.

The lack of precision as to why one wants theory seems to come from a fundamental dichotomy about what one wants of it, between (a) a human-comprehensible account that helps one think about some systems and (b) a formal account that empirically matches a range of observed cases. The ultimate, physics-inspired, hope seems to be that one would find a theory of socio-ecological systems that does both. The basic problem is we have not come up with something that is both, so that we either have abstract models that give us nice stories of emergence etc. But do not have an empirically rigorous relation to observed cases or models that relate closely to data from observed cases but not to a wide range of them.

There is no problem with developing ways of thinking about systems using abstract models as long as one does not delude oneself that this is an empirical relationship, but rather you are using the model to understand a set of *ideas* – similar to the analogical or theoretical exploration purposes in Edmonds et al. (2019). The difficulty comes if one wants theory to be empirical, e.g. to explain or predict observed data. One may hope that an abstract model might later be developed to become rigorously empirical, but one should only claim that once this has been established.

In terms of testing theory empirically, in the examples we found, the test was relative to a raft of other assumptions (in addition to those from the theory being tested). These assumptions were often not fully described. In particular, it was frequently not clear how essential or reliable these auxiliary assumptions were. If we are to use ABM to rigorously test theory we need at least two steps: (1) a more formal listing of all the assumptions in a model, describing each in terms of their provenance and reliability (e.g. A is empirically established, B is suggested by the literature, C is a traditional assumption, D seems to us based on common sense, E was necessary to get the simulation to work) and then (2) to systematically vary the relatively unreliable assumptions to check if the result of the testing remains the same. If we did this, we would know the reduced set of auxiliary assumptions under which the target theory held.

A similar approach with assumptions might be necessary in terms of inducing new empirically-valid theory, since we need to distinguish what is essential to the theory and what is irrelevant detail. Two approaches seem possible here.

1. The first is doing a lot of empirical models of specific cases of a certain class of system and then analysing what was common to all of those models. This might be aided by machine algorithms to compare mechanisms and outcomes automatically, but you still need the set of base case models to work on. This is a lot of work, so has not been done for socio-ecological systems as far as we can tell. It is possible that, by agreeing on some certain commonalities first (e.g. ontologies) that the subsequent comparison process would be easier.
2. The second is via a process of meta-modelling, starting with an empirical model and then modelling that model with a simpler one (Lafuerza et al., 2016a). In this approach the important outcomes are identified and then ensuring that any simplifications in the meta-model do not result in any significantly different outcomes. Thus, even if an initial meta-model over-simplifies (resulting in significantly different outcomes) more of the detail from the original can be reinstated until the align. This process can progress further to a meta-meta-model etc. (as in Lafuerza et al., 2016b).

The trouble with empirical science is that it takes a lot of work and the more complex the target system, the more work it takes. The amount of work probably means that such efforts need to be collectively, rather than individually, organised. It is far easier to deal with abstract models of ideas, but those can merely give the impression of progress. Furthermore, any theories that result may well not be simple, for simplicity is not a reliable indicator of empirical truth (Edmonds 2007).

8. Concluding summary

Equating maths with theory does not seem very helpful and risks confusing debate about which formal tool is most suitable for the task in hand with how to proceed in terms of generalisation. We recommend that this usage is avoided, since it does not increase clarity or enable useful distinction.

The use of theories as input is difficult due to (a) their vagueness (and hence many other assumptions are needed to instantiate these into an ABM) and (b) their reliability, since the theories often do not make clear their empirically established conditions of application (or they are not even known). Exploring *all* the possible additional assumptions to turn theory into an ABM (all or part of it) is usually infeasible. To get a new theory out of ABMs with any level of generality would require the comparison and analysis of several ABMs, which takes a lot of work and is rarely done. Any conclusions about theories put into an ABM is relative to a host of other added assumptions. Precision is needed in describing this kind of activity, so that readers are clear about how theory is used.

Whilst existing theories are often used to help specify ABMs, the link from what comes out of ABMs to theories as they are in their domain subjects is weak. We suggest that: (a) the link from model back to theory needs to be strengthened by collaboration with theory colleagues after simulation results are known, (b) more rigorous methodologies (maybe new inductive approaches such as sketched above) are needed if new theory is to be identified from ABMs, (c) the rhetoric about testing or developing new theory should be moderated in line with what can actually be achieved, maybe aiming for *some* level of generalisation rather than trying to 'jump' to supposedly very general conclusions (d) if we are serious about obtaining greater empirical generality for our models then we need to organise so we can collectively achieve this, and (e) decide what is important to us in terms of theory – what we want it *for*.

Declaration of competing interest

There are no conflicts of interest that we are aware of.

Data availability

No data was used for the research described in the article.

Acknowledgements

Edmonds was supported as part of the ESRC-funded, UK part of the "ToRealSim" project, grant number ES/S015159/1. Heppenstall was also supported by grants from UKPRP (MR/S037578/2), Medical Research Council (MC_UU_00022/5) and Scottish Government Chief Scientist Office (SPHSU20). Polhill was funded by the Scottish Government's Environment Natural Resources and Agriculture Strategic Research Programme 2022–2027 (JHI-C5-1). Wijermans was funded by the Swedish Research Council Formas, Project Grant No. 2018-00401. Thanks to all those in the workshops on Agents for Theory organised in Hannover for their valuable discussions and, in particular, to the organisers: Iris Lorscheid, Uta Berger, Volker Grimm and Matthias Meyer.

References

Ajzen, I., 1985. From intentions to actions: a theory of planned behavior. In: Kuhl, J., Beckmann, J. (Eds.), Action Control. Springer, pp. 11–39. https://doi.org/10.1007/978-3-642-69746-3_2 (Cited in Taghikhah et al. 2021).
Ajzen, I., 1991. The theory of planned behavior. Organ. Behav. Hum. Decis. Process. 50 (2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-t (Cited in Vu et al. 2020.
An, L., Grimm, V., Sullivan, A., Turner, B.L., Malleson, N., Heppenstall, A., Vincenot, C., Robinson, D., Ye, X., Liu, J., Lindkvist, E., Tang, W., 2021. Challenges, tasks, and opportunities in modeling agent-based complex systems. Ecol. Model. 457, 109685.
Anderson, P., 1999. Perspective: complexity theory and organization science. Organ. Sci. 10 (3), 216–232.
Armano, S., Drazen, P., Petra, R., Kruno, K., 2003. An agent-based model of ethnic mobilisation. J. Artif. Soc. Soc. Simulat. 6 (1), 1. https://www.jasss.org/6/1/1.html.
Balzer, W., Brendel, K.R., Hofmann, S., 2001. Bad arguments in the comparison of game theory and simulation in social studies. J. Artif. Soc. Soc. Simulat. 4 (2), 1. https://www.jasss.org/4/2/1.html.
Barnaud, C., Le Page, C., Dumrongrojwatthana, P., Trébuil, G., 2013. Spatial representations are not neutral: lessons from a participatory agent-based modelling process in a land-use conflict. Environ. Model. Software 45, 150–159.
Becker, G.S., Murphy, K.M., 1988. A theory of rational addiction. J. Polit. Econ. 96 (4), 675–700 (Cited in Vu et al. 2020).
Bergman, N., Haxeltine, A., Whitmarsh, L., Kohler, J., Schilperoord, M., Rotmans, J., 2008. Modelling socio-technical transition patterns and pathways. J. Artif. Soc. Soc. Simulat. 11 (3), 7. https://www.jasss.org/11/3/7.html.
Bonabeau, E., 2002. Predicting the unpredictable. Harv. Bus. Rev. 80 (3), 109–116.
Braithwaite, R.B., 1962. Models in the empirical sciences. In: Nagel, Suppes, and Tarski 1962, pp. 224–231.
Carley, K.M., 1995. Computational and mathematical organization theory: perspective and directions. Comput. Math. Organ. Theor. 1 (1), 39–56.
Christaller, W., 1933. Die Zentralen Orte in Süddeutschland. Fischer, Jena.
Cialdini, R.B., Kallgren, C.A., Reno, R.R., 1991. A focus theory of normative conduct: a theoretical refinement of the role of norms in human behavior. In: Zanna, M.P. (Ed.), Advances in Experimental Social Psychology, vol. 24. Academic Press, pp. 201–234. https://doi.org/10.1016/s0065-2601(08)60330-5 (Cited in Vu et al. 2020).
Coelho, C.G.C., Ralha, C.G., 2022. MASE-EGTI: an agent-based simulator for environmental land change. Environ. Model. Software 147, 105252. https://doi.org/10.1016/j.envsoft.2021.105252.
Coleman, J.S., 1986. Social theory, social research, and a theory of action. Am. J. Sociol. 91 (6), 1309–1335. https://doi.org/10.1086/228423 (Cited in Spaiser and Sumpter 2016; Vu et al. 2020).
Crooks, A., 2010. Constructing and implementing an agent-based model of residential segregation through vector GIS. IJGIS 24 (5), 661–675.
Crooks, A., Castle, C., Batty, M., 2008. Key challenges in agent-based modelling for geo-spatial simulation. Comput. Environ. Urban Syst. 32, 417–430.
Dilaver, O., 2015. From participants to agents: grounded simulation as a mixed-methods research design. J. Artif. Soc. Soc. Simulat. 18 (1), 15. https://www.jasss.org/18/1/15.html.
Edmonds, B., 2007. Simplicity is not truth-indicative. In: Gershenson, C. (Ed.), (2007) Philosophy and Complexity. World Scientific, pp. 65–80.
Edmonds, B., le Page, C., Bithell, M., Chattoe-Brown, E., Grimm, V., Meyer, R., Montañola-Sales, C., Ormerod, P., Root, H., Squazzoni, F., 2019. Different modelling

purposes. J. Artif. Soc. Soc. Simulat. 22 (3), 6. http://jasss.soc.surrey.ac.uk/22/3/6.html.
Elsenbroich, C., 2012. Explanation in agent-based modelling: functions, causality or mechanisms? J. Artif. Soc. Soc. Simulat. 15 (3), 1. http://jasss.soc.surrey.ac.uk/15/3/1.html.
Epstein, J.M., 2008. Why model? J. Artif. Soc. Soc. Simulat. 11 (4), 12. https://jasss.soc.surrey.ac.uk/11/4/12.html.
Festinger, L., 1962. A Theory of Cognitive Dissonance. Stanford University Press (Cited in Taghikhah et al. 2021).
Filatova, T., 2015. Empirical agent-based land market: Integrating adaptive economic behavior in urban land-use models. Comput. Environ. Urban Syst. 54, 397–413.
Fonoberova, M., Fonoberov, V.A., Mezic, I., Mezic, J., Brantingham, J., P., 2012. Nonlinear dynamics of crime and violence in urban settings. J. Artif. Soc. Soc. Simulat. 15 (1), 2. http://jasss.soc.surrey.ac.uk/15/1/2.html.
Galan, J.M., Izquierdo, L.R., 2005. Appearances can Be deceiving: lessons learned Re-implementing axelrod's 'evolutionary approach to norms. J. Artif. Soc. Soc. Simulat. 8 (3), 2. https://www.jasss.org/8/3/2.html.
Giere, R.N., 2010. Explaining Science. University of Chicago Press.
Gilbert, N., Pyka, A., Ahrweiler, P., 2001. Innovation networks-a simulation approach. J. Artif. Soc. Soc. Simulat. 4 (3), 1–13. https://www.jasss.org/4/3/8.html.
Glaser, B.G., Strauss, A.L., 1967. The Discovery of Grounded Theory: Strategies for Qualitative Research. Aldine, Chicago (Cited in Dilaver 2015).
Groeneveld, J., Müller, B., Buchmann, C.M., Dressler, G., Guo, C., Hase, N., Hoffmann, F., John, F., Klassert, C., Lauf, T., Liebelt, V., Nolzen, H., Pannicke, N., Schulze, J., Weise, H., Schwarz, N., 2017. Theoretical foundations of human decision-making in agent-based land use models – a review. Environ. Model. Software 87, 39–48.
Hägerstrand, T., 1952. The Propagation of Innovation Waves. Lund Studies in Geography: Series B, Human Geography, vol. 4. Royal University of Lund, Dept. of Geography, Lund. OCLC 254752.
Hartshorn, M., Kaznatcheev, A., Shultz, T., 2013. The evolutionary dominance of ethnocentric cooperation. J. Artif. Soc. Soc. Simulat. 16 (3), 7. http://jasss.soc.surrey.ac.uk/16/3/7.html.
Hassan, S., Arroyo, J., Galán, J.M., Antunes, L., Pavón, J., 2013. Asking the Oracle: introducing forecasting principles into agent-based modelling. J. Artif. Soc. Soc. Simulat. 16 (3), 13. https://jasss.soc.surrey.ac.uk/16/3/13.html.
Hedström, P., Swedberg, R., 1996. Social mechanisms. Acta Sociol. 39 (3), 281–308. https://doi.org/10.1177/000169939603900302.
Hedström, P., Ylikoski, P., 2010. Causal mechanisms in the social sciences. Annu. Rev. Sociol. 36, 49–67.
Helbing, D., Balietti, S., 2011. How to do agent-based simulations in the future: from modeling social mechanisms to emergent phenomena and interactive systems design. In: Technical Report, Santa Fe Institute, Working Paper 11-06-024.
Hempel, C.G., 1942. The function of general laws in history. J. Philos. 39 (2), 35–48. https://doi.org/10.2307/2017635 (reprinted with slight modifications in Hempel 1965: 231–244.
Hempel, C.G., 1965. Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. Free Press, New York.
Hempel, C.G., 1974. Formulation and formalization of scientific theories. In: Suppe, F. (Ed.), The Structure of Scientific Theories. University of Illinois Press, Urbana, pp. 244–254.
Hoertel, N., Blachier, M., Blanco, C., Olfson, M., Massetti, M., Rico, M.S., Limosin, F., Leleu, H., 2020. A stochastic agent-based model of the SARS-CoV-2 epidemic in France. Nat. Med. 26 (9), 1417–1421.
Holtz, G., 2014. Generating social practices. J. Artif. Soc. Soc. Simulat. 17 (1), 17. http://jasss.soc.surrey.ac.uk/17/1/17.html.
Huang, Q., Parker, D.C., Filatova, T., Sun, S., 2013. A review of urban residential choice models using agent-based models. Environ. Plann. B 40.
Ingelhart, R., Welzel, C., 2005. Modernization, Cultural Change and Democracy: the Human Development Sequence. Cambridge University Press, Cambridge (Cited in Spaiser and Sumpter 2016).
Ivanova, O., Kancs, A., Thissen, M., 2019. EU Economic Modelling System: Assessment of the European Institute of Innovation and Technology (EIT) Investments in Innovation and Human Capital. https://doi.org/10.2791/184008. Luxembourg.
Izquierdo, S.S., Izquierdo, L.R., Gotts, N.M., 2008. Reinforcement learning dynamics in social dilemmas. J. Artif. Soc. Soc. Simulat. 11 (2), 1. https://www.jasss.org/11/2/1.html.
Izquierdo, L.R., Izquierdo, S.S., Galán, J.M., Santos, J.I., 2009. Techniques to understand computer simulations: Markov chain analysis. J. Artif. Soc. Soc. Simulat. 12 (1), 6. https://www.jasss.org/12/1/6.html.
Jager, W., Janssen, M., 2012. An updated conceptual framework for integrated modeling of human decision making: the Consumat II. In: Paper for Workshop Complexity in the Real World@ ECCS, pp. 1–18.
Jager, W., Janssen, M.A., De Vries, H.J.M., De Greef, J., Vlek, C.A.J., 2000. Behaviour in commons dilemmas: Homo economicus and Homo psychologicus in an ecological economic model. Ecol. Econ. 35, 357–379.
Knibbe, R., Drop, M., Muytjens, A., 1987. Correlates of stages in the progression from everyday drinking to problem drinking. Soc. Sci. Med. 24 (5), 463–473. https://doi.org/10.1016/0277-9536(87)90221-8 (Cited in Vu et al. 2020).
Knoeri, C., Binder, C.R., Althaus, H.J., 2011. An agent operationalization approach for context specific agent-based modeling. J. Artif. Soc. Soc. Simulat. 14 (2), 4. https://jasss.soc.surrey.ac.uk/14/2/4.html.
Koomen, E., Diogo, V., Dekkers, J., Rietveld, P., 2015. A utility-based suitability framework for integrated local-scale land-use modelling. Comput. Environ. Urban Syst. 50, 1–14.
Lafuerza, L.F., Dyson, L., Edmonds, B., McKane, A.J., 2016a. Staged models for interdisciplinary research. PLoS One 11 (6), e0157261. https://doi.org/10.1371/journal.pone.0157261.
Lafuerza, L.F., Dyson, L., Edmonds, B., McKane, A.J., 2016b. Simplification and analysis of a model of social interaction in voting. Eur. Phys. J. B 89, 159. https://doi.org/10.1140/epjb/e2016-70062-2.
Li, J., Xiao, R., 2017. Agent-based modelling approach for multidimensional opinion polarization in collective behaviour. J. Artif. Soc. Soc. Simulat. 20 (2), 4. https://jasss.soc.surrey.ac.uk/20/2/4.html.
Lindenberg, S., Steg, L., 2011. Normative, gain and hedonic goal frames guiding environmental behavior. J. Soc. Issues 63 (2), 117–137. https://doi.org/10.1111/j.1540-4560.2007.00499.x.
Liu, C., O'Sullivan, D., 2016. An abstract model of gentrification as a spatially contagious succession process. Comput. Environ. Urban Syst. 59, 1–10.
Lorscheid, I., Berger, U., Grimm, V., Meyer, M., 2019. From cases to general principles: a call for theory development through agent-based modeling. Ecol. Model. 393, 153–156.
Lustick, I.S., 2000. Agent-based modelling of collective identity: testing constructivist theory. J. Artif. Soc. Soc. Simulat. 3 (1), 1. https://www.jasss.org/3/1/1.html.
Macy, M.W., Willer, R., 2002. From factors to actors: computational sociology and agent-based modeling. Annu. Rev. Sociol. 28 (1), 143–166.
Magliocca, N., Safirova, E., McConnell, V., Walls, M., 2011. An economic agent-based model of coupled housing and land markets (CHALMS). Comput. Environ. Urban Syst. 35 (3), 183–191.
Moss, S., 2008. Alternative approaches to the empirical validation of agent-based models. J. Artif. Soc. Soc. Simulat. 11 (1), 5. https://jasss.soc.surrey.ac.uk/11/1/5.html.
Muelder, H., Filatova, T., 2018. One theory - many formalizations: testing different code implementations of the theory of planned behaviour in energy agent-based models. J. Artif. Soc. Soc. Simulat. 21 (4), 5. http://jasss.soc.surrey.ac.uk/21/4/5.html.
Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., Schwarz, N., 2013. Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. Environ. Model. Software 48, 37–48.
Nagel, E., Suppes, P., Tarski, A., 1962. Logic, methodology and the philosophy of science. In: Proceedings of the 1960 International Congress. Stanford University Press, Stanford.
Niamir, L., Ivanova, O., Filatova, T., 2020. Economy-wide impacts of behavioral climate change mitigation: linking agent-based and computable general equilibrium models. Environ. Model. Software 134, 104839. https://doi.org/10.1016/j.envsoft.2020.104839.
O'Sullivan, D., Evans, T., Manson, S., Metcalf, S., Ligmann-Zielinska, A., Bone, C., 2016. Strategic directions for agent-based modeling: avoiding the YAAWN syndrome. J. Land Use Sci. 11, 177–187.
Pacilly, F., Hofstede, G.J., Lammerts van Bueren, E., Groot, J., 2019. Analysing social-ecological interactions in disease control: an agent-based model on farmers' decision making and potato late blight dynamics. Environ. Model. Software 119, 354–373.
Parker, D.C., Meretsky, V., 2004. Measuring pattern outcomes in an agent-based model of edge-effect externalities using spatial metrics. Agric. Ecosyst. Environ. 101 (2–3), 233–250.
Picascia, S., 2017. Agent-based Modelling of Urban Economic and Cultural Dynamics under the Rent-Gap Hypotheses. Doctoral dissertation, Manchester Metropolitan University. https://e-space.mmu.ac.uk/618397/1/thesis-final.pdf.
Power, C., 2009. A spatial agent-based model of N-person prisoner's dilemma cooperation in a socio-geographic community. J. Artif. Soc. Soc. Simulat. 12 (1), 8. https://www.jasss.org/12/1/8.html.
Rand, W., Herrmann, J., Schein, B., Vodopivec, N., 2015. An agent-based model of urgent diffusion in social media. J. Artif. Soc. Soc. Simulat. 18 (2), 1. http://jasss.soc.surrey.ac.uk/18/2/1.html.
Rauhut, H., Junker, M., 2009. Punishment deters crime because humans are bounded in their strategic decision-making. J. Artif. Soc. Soc. Simulat. 12 (3), 1. https://www.jasss.org/12/3/1.html.
Rimal, R.N., Real, K., 2005. How behaviors are influenced by perceived norms: a test of the Theory of Normative Social Behavior. Commun. Res. 32 (3), 389–414. https://doi.org/10.1177/0093650205275385 (Cited in Vu et al. 2020).
Robinson, D.T., Murray-Rust, D., Rieser, V., Milicic, V., Rounsevell, M., 2012. Modelling the impacts of land system dynamics on human well-being: Using an agent-based approach to cope with data limitations in Koper, Slovenia. Comput. Environ. Urban Syst. 36 (2), 164–176.
Rosés, R., Kadar, C., Malleson, N., 2021. A data-driven agent-based simulation to predict crime patterns in an urban environment. Comput. Environ. Urban Syst. 89, 101660.
Sakoda, J.M., 1971. The checkerboard model of social interaction. J. Math. Sociol. 1 (1), 119–132.
Salzarulo, L., 2006. A continuous opinion dynamics model based on the principle of meta-contrast. J. Artif. Soc. Soc. Simulat. 9 (1), 13. https://www.jasss.org/9/1/13.html.
Sasaki, Y., Box, P., 2003. Agent-based verification of von Thünen's location theory. J. Artif. Soc. Soc. Simulat. 6 (2), 9. https://www.jasss.org/6/2/9.html.
Sawyer, R.K., 2005. Social Emergence: Societies as Complex Systems. Cambridge University Press.
Schelling, T.C., 1971. Dynamic models of segregation. J. Math. Sociol. 1 (1), 143–186.
Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., et al., 2017. A framework for mapping and comparing behavioural theories in models of social-ecological systems. Ecol. Econ. 131, 21–35.

Schulze, J., Müller, B., Groeneveld, J., Grimm, V., 2017. Agent-based modelling of social-ecological systems: achievements, challenges, and a way forward. J. Artif. Soc. Soc. Simulat. 20 (2) 8, 2017. https://jasss.soc.surrey.ac.uk/20/2/8.html.
Schwartz, S.H., 1977. Normative influence on altruism. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology, vol. 10. Academic Press, New York, pp. 221–279 (1977).
Smajgl, A., Barreteau, O. (Eds.), 2014. Empirical Agent-Based Modelling - Challenges and Solutions. Springer, New York, NY.
Smaldino, P., Pickett, C., Sherman, J., Schank, J., 2012. An agent-based model of social identity dynamics. J. Artif. Soc. Soc. Simulat. 15 (4), 7. http://jasss.soc.surrey.ac.uk/15/4/7.html.
Spaiser, V., Sumpter, D.J.T., 2016. Revising the Human Development Sequence theory using and agent-based approach and data. J. Artif. Soc. Soc. Simulat. 19 (3), 1. https://jasss.soc.surrey.ac.uk/19/3/1.html.
Staller, A., Petta, P., 2001. Introducing emotions into the computational study of social norms: a first evaluation. J. Artif. Soc. Soc. Simulat. 4 (1), 2. https://www.jasss.org/4/1/2.html.
Stern, P., Dietz, T., Abel, T., Guagnano, G., Kalof, L., 1999. A value-belief-norm theory of support for social movements: the case of environmentalism. Hum. Ecol. Rev. 6.
Sun, Z., Lorscheid, I., Millington, J.D., Lauf, S., Magliocca, N.R., Groeneveld, J., Balbi, S., Nolzen, H., Müller, B., Schulze, J., Buchman, C.M., 2016. Simple or complicated agent-based models? A complicated issue. Environ. Model. Software 86, 56–67. https://doi.org/10.1016/j.envsoft.2016.09.006 (Cited in Taghikhah et al. 2021.
Suppe, F., 2000. Understanding scientific theories: An assessment of developments, 1969–1998. Philos. Sci. 67 (S3), S102–S115.
Taghikhah, F., Filatova, T., Voinov, A., 2021. Where does theory have it right? A comparison of theory-driven and empirical agent based models. J. Artif. Soc. Soc. Simulat. 24 (2), 4. https://jasss.soc.surrey.ac.uk/24/2/4.html.
Tarski, A., 1936. On the concept of logical consequence. Logic, semantics, metamathematics 52, 409–420.
Thiele, J.C., Kurth, W., Grimm, V., 2014. Facilitating parameter estimation and sensitivity analysis of agent-based models: a cookbook using NetLogo and 'R. J. Artif. Soc. Soc. Simulat. 17 (3), 11. http://jasss.soc.surrey.ac.uk/17/3/11.html.
Tobler, W., 1970. A computer movie simulating urban growth in the Detroit region. Econ. Geogr. 46 (Suppl. ment), 234–240.
Torrens, P.M., Nara, A., 2007. Modeling gentrification dynamics: A hybrid approach. Comput. Environ. Urban Syst. 31 (3), 337–361.
Touza, J., Drechsler, M., Smart, J., Termansen, M., 2013. Emergence of cooperative behaviours in the management of mobile ecological resources. Environ. Model. Software 45, 52–63.
Troitzsch, K.G., 2009. 'Not all explanations predict satisfactorily, and not all good predictions explain'. J. Artif. Soc. Soc. Simulat. 12 (1), 10. https://www.jasss.org/12/1/10.html.
Vallacher, R.R., Nowak, A., 1994. The chaos in social psychology. In: Vallacher, R.R., Nowak, A. (Eds.), Dynamical Systems in Social Psychology. Academic Press, San Diego, pp. 1–6.
Zhu, H., Wang, F., 2021. An agent-based model for simulating urban crime with improved daily routines. Comput. Environ. Urban Syst. 89, 101680.
Zhuge, C., Shao, C., Gao, J., Dong, C., Zhang, H., 2016. Agent-based joint model of residential location choice and real estate price for land use and transport model. Comput. Environ. Urban Syst. 57, 93–105.
van der Kam, M., Peters, A., van Sark, Wilfried, Floor, A., 2019. Agent-based modelling of charging behaviour of electric vehicle drivers. J. Artif. Soc. Soc. Simulat. 22 (4), 7. http://jasss.soc.surrey.ac.uk/22/4/7.html.
Voinov, A., Shugart, H.H., 2013. 'Integronsters', integral and integrated modeling. Environ. Model. Software 39, 149–158.
Vu, T.M., Probst, C., Nielsen, A., Bai, H., Meier, P.S., Buckley, C., Strong, M., Brennan, A., Purshouse, R.C., 2020. A software architecture for mechanism-based social systems modelling in agent-based simulation models. J. Artif. Soc. Soc. Simulat. 23 (3), 1. https://www.jasss.org/23/3/1.html.
Wartofsky, M.W., 1979. The Model Muddle: Proposals for an Immodest Realism: [1966]. In Models: Representation and the scientific understanding. Springer Netherlands, Dordrecht, pp. 1–11.
Wijermans, N., Jorna, R., Jager, W., van Vliet, T., Adang, O., 2013. CROSS: modelling crowd behaviour with social-cognitive agents. J. Artif. Soc. Soc. Simulat. 16 (4), 1. http://jasss.soc.surrey.ac.uk/16/4/1.html.
Wildman, W.J., Sosis, Ri., 2011. Stability of groups with costly beliefs and practices. J. Artif. Soc. Soc. Simulat. 14 (3), 6. http://jasss.soc.surrey.ac.uk/14/3/6.html.
Wilsnack, R.W., Cheloha, R., 1987. Women's roles and problem drinking across the lifespan. Soc. Probl. 34 (3), 231–248. https://doi.org/10.1525/sp.1987.34.3.03a00020 (Cited in Vu et al. 2020).
Wittgenstein, L., 1953. Philosophical Investigations. Blackwell Publishing.
Wu, F., 2002. Complexity and urban simulation: towards a computational laboratory. Geogr. Res. Forum 22, 22–40.
Yamaguchi, K., Kandel, D.B., 1985. On the resolution of role incompatibility: a life event history analysis of family roles and marijuana use. Am. J. Sociol. 90 (6), 1284–1325. https://doi.org/10.1086/228211 (Cited in Vu et al. 2020.
Zepeda, L., Deal, D., 2009. Organic and local food consumer behaviour: alphabet theory. Int. J. Consum. Stud. 33 (6), 697–705. https://doi.org/10.1111/j.1470-6431.2009.00814.x (Cited in Taghikhah et al. 2021.