

# Formalising theories of human decision-making for agent-based modelling of social-ecological systems: practical lessons learned and ways forward

Nina Schwarz<sup>1,2\*</sup>, Gunnar Dressler<sup>3</sup>, Karin Frank<sup>3</sup>, Wander Jager<sup>4</sup>, Marco Janssen<sup>5</sup>, Birgit Müller<sup>3</sup>, Maja Schlüter<sup>6</sup>, Nanda Wijermans<sup>6</sup>, Jürgen Groeneveld<sup>3,7,8</sup>

<sup>1</sup>University of Twente, faculty of Geo-Information Science and Earth Observation, Enschede, The Netherlands

<sup>2</sup>UFZ-Helmholtz Centre for Environmental Research, Department of Computational Landscape Ecology, Leipzig, Germany

<sup>3</sup>UFZ-Helmholtz Centre for Environmental Research, Department of Ecological Modelling, Leipzig, Germany

<sup>4</sup>University College Groningen, Groningen, The Netherlands

<sup>5</sup>School of Sustainability, Arizona State University, Tempe, Arizona, USA

<sup>6</sup>Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden

<sup>7</sup>Department of Economics, Leipzig University and German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany.

<sup>8</sup>Institute of Forest Growth and Forest Computer Sciences, Technische Universität Dresden, Tharandt, Germany

---

---

## Abstract

Incorporating representations of human decision-making that are based on social science theories into social-ecological models is considered increasingly important – yet choosing and formalising a theory for a particular modelling context remains challenging. Here, we reflect on our experiences of selecting, formalising and documenting psychological and economic theories of human decision-making for inclusion in different agent-based models (ABMs) of natural resource use. We discuss the challenges related to four critical tasks: How to select a theory? How to formalise a theory and how to translate it into code? How to document the formalisation? In this way, we present a systematic overview of the choices researchers face when including theories of human decision-making in their ABMs, reflect on the choices we made in our own modelling projects and provide guidance for those new to the field. Also, we highlight further challenges regarding the parameterisation and analysis of such ABMs and suggest that a systematic overview of how to tackle these challenges contributes to an effective collaboration in interdisciplinary teams addressing socio-ecological dynamics using models.

## Keywords

Agent-based modelling; theories on human decision-making; model documentation; formalisation; natural resource use

## Code availability

Three models are discussed in this paper using the open Agent-Based modelling software NetLogo (<https://ccl.northwestern.edu/netlogo/>). Code source and documentation, along with information to analyse and replicate the results, are available for Daisyworld at <https://www.comses.net/codebases/4958/releases/1.0.0/>, RAGE at <https://www.comses.net/codebases/5721/releases/1.0.6/>, and FIBE at <https://www.comses.net/codebases/4d9cfac5-7331-4a03-9d83-ab06cbedc143/releases/1.0.0/>.

---

---

## Correspondence:

Contact N. Schwarz at [n.schwarz@utwente.nl](mailto:n.schwarz@utwente.nl)

## Cite this article as:

Schwarz, N., Dressler, G., Frank, K., Jager, W., Janssen, M., Müller, B., Schlüter, M., Wijermans, N., & Groeneveld, J.

Formalising theories of human decision-making for agent-based modelling of social-ecological systems: practical lessons learned and ways forward

*Socio-Environmental Systems Modelling*, vol. 2, 16340, 2020, doi:10.18174/sesmo.2020a16340

This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/).



**Socio-Environmental Systems Modelling**

An Open-Access Scholarly Journal

<http://www.sesmo.org>

## 1. Introduction

Incorporating representations of human behaviour in social-ecological models contributes to a better understanding of the often complex interactions between humans and their natural environment, and it offers possibilities for experimentation with strategies or policies aimed at changing behaviour (Jager & Ernst, 2017; Schulze et al., 2017). Agent-based models (ABMs) allow for incorporating human behaviour into a model since they enable representing individual agents, their decision-making and the interactions among themselves and with their environment.

Humans often make decisions following norms, habits, heuristics and different learning strategies, among other factors (e.g. World Bank, 2015), and such processes are also included in simulation models (Heckbert et al., 2010; Yu et al., 2019). In particular, decision strategies that employ social influences (e.g. norms) may trigger processes of social diffusion, i.e., the spread of innovations in a social system. Such processes may result in tipping points in social systems (Nyborg et al., 2016), and are the subject of theories on human decision-making developed in the social sciences. Here, we focus in particular on decision-making of individuals, and particularly on theories from behavioural economics, social and environmental psychology, which are commonly used to model individual decision-making in stable social and ecological contexts. Not only the social context influences human decision-making and behaviour: There is increasing evidence that human decision-making and behaviour in natural resource contexts is also influenced by the biophysical conditions in a particular place (Masterson et al., 2017; Muhar et al., 2018).

Following the modelling perspective taken in Groeneveld et al. (2017), a specific theory is rooted either in

- rationality (i.e. actors with consistent and well-defined preferences, choosing options best meeting their preferences and taking into account all relevant information),
- bounded rationality (i.e. deviating from one or more of the assumptions of rationality; Levine et al., 2015; Simon, 1955), or
- non-rationality (i.e. actors do not actively pursue any explicit or implicit goal in their decision-making but, for instance, choose randomly).

Although many theories on human decision-making related to these three backgrounds or meta-theories exist, their use in social-ecological modelling has been very limited to date (Groeneveld et al., 2017). Instead, modellers, for instance working on coupled human-natural systems (An, 2012) or land use/cover change (Groeneveld et al., 2017) often develop ad-hoc implementations of human behaviour to represent the decision processes observed in the field, or they base the decision model on meta-theories of the (bounded) rational actor.

Depending on the model context and purpose, ad-hoc implementations of human decision-making may be perfectly reasonable. For instance, engaging stakeholders in participatory modelling where they set up the agents' rules is an obvious model purpose that almost prohibits using a theory. In addition, big data together with data mining methods predicting human decision-making does not require any underlying theory of individual decision making. It has been even discussed that such methods may make the use of theories obsolete (Anderson, 2008). We do not discuss machine learning and big data approaches here, since our focus is on models that try to increase our understanding how actors behave from an individual perspective. For many model contexts and purposes, building ABMs on existing theories on human decision-making can have manifold benefits. First, informing a model with a behavioural theory fosters (a) re-using and comparing models since these theories could serve as a common denominator. Reusing and comparing models can save time and accelerate scientific advancements as well as foster the development of theory (Bell et al., 2015; Crooks et al., 2008). (b) Theory-based models can be used to test alternative theories with empirical data (Klabunde & Willekens, 2016; Silverman et al., 2011). This can inform theory development, since theories are not always well-established and backed up by empirical data, and modelling may point out inadequacies (Davis et al., 2018). For instance, many psychological and economic theories on human decision-making are based on drivers of decision-making and end with the decision or performance of the behaviour. Therefore, how the behaviour of many individuals (macro level) shapes the social and ecological setting that determines the decision-making of an individual (micro-level) in a subsequent stage is not addressed. ABMs are a useful tool to study these micro-macro interactions as they simulate macro-level properties and behaviours that influence future choices of individual agents. Second, specifically considering psychological and economic theories of human decision-making can have additional benefits: (c) Incorporating such theories into the decision process fosters modelling

of behavioural drivers and processes that may yield insights into the possible societal responses to different events and policy measures or changes in the biophysical environment. (d) At the same time, to base an agent's decision-making on a theory can help narrow down the seemingly endless options of what could be included in the model to those processes that are deemed relevant by the theory (Edmonds, 2017). (e) Finally, this approach may provide a common framing in interdisciplinary teams and facilitate communication between modellers and social scientists not yet familiar with the other domain (Davis et al., 2007).

Not surprisingly, the importance of incorporating representations of human decision-making that are based on social science theories is also being increasingly acknowledged for social-ecological models (Elsawah et al., 2020). Choosing a relevant theory/theories of human decision-making and formalising them remains difficult (see, e.g., Jager, 2017; Poile & Safayeni, 2016). There are several reasons why explicitly including theories in a model is challenging (Davis et al., 2018; Schlüter et al., 2017). The largest hurdle is the more descriptive nature of most theories on human decision-making, which are presented as narratives: Theories provide far less detail than what is required for a simulation model (Poile & Safayeni, 2016) or sometimes even to derive testable hypotheses from them (Sullivan, 2002) and compare theories (West et al., 2019). As a consequence, different formalisations of the same theory are possible, as demonstrated for the Theory of Planned Behaviour by Muelder & Filatova (2018). Second, the modeller is confronted with an abundance of different theories on human decision-making that often overlap and apply to different levels of aggregation, which makes it difficult to a) obtain an overview of the wide field of theories that address human decision-making and b) find theories that fit a specific context. Third, psychological theories are often expressed as correlations between theoretical constructs (Jager, 2017). When formalising such theories in computational models, these correlations first need to be translated into causal relationships allowing for simulation.

Several steps have already been taken to facilitate the inclusion of theories of human decision-making, specifically in ABMs. First, literature reviews have compiled theories that have been used in models (An, 2012; Cooke et al., 2009; Groeneveld et al., 2017; Jager, 2017; Klabunde & Willekens, 2016) as well as those that could potentially be useful in models (Brenner, 2006; Meyfroidt, 2013). Second, the *MoHuB* (*Modelling Human Behaviour*) framework supports the formalisation of theories on human decision-making of a natural resource user in social-ecological models by providing a common language to describe, compare and communicate alternative theories (Schlüter et al., 2017). West et al. (2019) pursue a similar purpose by providing a formal system for describing theories based on a library of constructs of its components and relationships between pairs of constructs. Third, Poile and Safayeni (2016) point out that, due to missing details in theories, many different assumptions can be taken to formalise a theory. Different combinations of these assumptions can lead to equifinality, i.e., identical outcomes stemming from different processes. The authors illustrate how this can be a problem with a simple simulation aiming at replicating a social psychological experiment. There is little in the literature that reviews concrete formalisations of theories into computer models, with the exception of Jager (2017) who provides a few detailed descriptions using example models. However, there is sufficient evidence that a specification problem exists, and a systematic overview of the theories and related assumptions that are incorporated when formalising human decision-making in an ABM is overdue.

This is where our paper steps in. We aim to encourage and guide modellers to incorporate theories of human decision-making into a computational model by sharing our experiences. We reflect on the practice of explicitly exploring human decision-making in models of social-ecological systems. By discussing those examples and reflecting on common challenges the reader learns about how to go from ambition to implementation. To follow up on the challenges described by Schlüter et al. (2017), we identified four tasks of i) finding psychological and economic theories about individuals taking decisions, ii) formalising and iii) implementing them and iv) documenting their implementation in ABMs. We demonstrate these tasks with models from three different resource use contexts (a fishery, a pastoral system and an abstract ecosystem services model for climate regulation) and highlight the possibilities, difficulties and choices one faces when incorporating a theory on human decision-making into a social-ecological ABM. We do not provide off-the-shelf solutions for formalising and implementing particular theories but rather demonstrate and reflect on the process from theory selection to formalisation, implementation and documentation. To do so, we propose a systematic set of tasks, more detailed questions as well as existing material we used that can help guide other modellers to incorporate theories of human decision-making into a computational model. These tasks refer to the model design and formulation and the implementation phases of the modelling cycle (Schmolke et al., 2010). For each task, we describe and reflect upon our experiences tackling them in the context of developing ABMs of natural resource use.

The remainder of the paper is organised as follows: section 2 introduces the three ABMs we are using as showcase models to describe our experiences and provides an overview of the theories used; section 3 shares our experiences regarding the four main challenges divided into four tasks for including theories of human decision-making into ABMs tackled here, namely, how to select a theory (section 3.1), how to formalise it (section 3.2), how to implement it (3.3) and how to document the formalisation (section 3.4). Section 4 points out issues related to theory implementation in later stages of the modelling cycle; and section 5 concludes.

## 2. Showcase models and theories to highlight our experiences

Throughout the paper, we often refer to specific theories of human decision-making, and all the theories used in our showcase models are summarised in Table 1. Table S1 in the Supplementary Material provides an overview of all other theories mentioned. Furthermore, when describing the decision-making process of an agent, we use the terminology introduced by Schlüter et al. (2017), namely:

- *State*: internal state of an individual
- *Behavioural options*: a set of options from which the individual can choose
- *Perception*: the process of sensing the surrounding social and biophysical environment
- *Evaluation*: the process of determining the significance, worth, or condition of the perceived state of the social and biophysical environment
- *Selection*: the process of choosing a behaviour from the set of perceived behavioural options
- *Behaviour*: the process of executing the chosen behavioural option

**Table 1:** Overview of the (meta-) theories used in the showcase models.

Theory	Main assumptions	Key references	Showcase models
Bounded rationality / satisficing	Bounded rationality: meta-theory assuming that rationality is limited by the available information and cognitive capacity	Gigerenzer and Selten (2001)	Daisyworld FIBE RAGE
	Satisficing: selection process that aims at reaching an aspirational level	Simon (1955)	
Descriptive norm / focus theory of normative conduct	Observing the behaviour of the majority or a reference group can impact a person's behaviour	Cialdini et al. (1990; 1991)	Daisyworld RAGE
Habitual behaviour/ reinforcement learning	Habits are behaviours that are often performed without much deliberation and can be formed by repeated performance of a rewarded behaviour (reinforcement learning)	Graybiel (2008), Pavlov (1927), Skinner (1938)	Daisyworld
Rational choice	Meta-theory assuming maximisation of self-interested utility with perfect knowledge and unlimited cognitive capacity	Frank (1987), Monroe (2001), Simon (1978)	Daisyworld RAGE
Social learning and imitation	Behaviour is learnt through observation and imitation of others	Bandura (1977)	FIBE
Theory of Planned Behaviour	Behaviour is explained by intentions and perceived behavioural control, i.e., beliefs about control over the behaviour. Intentions are based on attitudes towards the behaviour, subjective norm (approval by important others and willingness to comply with them), and perceived behavioural control	Ajzen (1991, 2012)	Daisyworld

Note: See Schlüter et al. (2017) for a more detailed overview of some of the theories listed here.

To explore and systematise the challenges of formalising theories of human decision-making, three ABMs were developed with different natural resource use contexts. These ABMs formalise a similar set of psychological and economic theories of human decision-making by first employing a core set of theories for all models and then complementing individual ABMs with additional theories (Table 2). Our recommendations draw upon our experiences building these three models as well as our experiences and insights from engaging in many other modelling exercises. These three ABM models investigate natural resource use in very different social-ecological contexts as follows.

The expanded Daisyworld model (Janssen, 2016) is a stylised model of resource use that uses the well-established Daisyworld model (Watson and Lovelock, 1983), in which black and white daisies provide ecosystem services for climate regulation. In this expanded model, agents were added to harvest daisies in their locality; the agents decide whether to move and whether to harvest daisies of which colour in each time step. The purpose of the Daisyworld model is to explore the effect of alternative theories of human decision-making on social-ecological dynamics in a generic setting.

The FIBE model (Wijermans et al., 2020) is a fishery model that represents the diversity of fisher behaviours as observed and classified in the (Swedish) Baltic Sea fishery. The purpose of the FIBE model is to explore the consequences of behavioural diversity for the sustainability and management of the fishery. The agents represent fishers, and the fish are the natural resource. The decision-making involves 1) whether the fishers will go fishing and, if so, 2) where they will fish. Fishers belong to one of three different fishing styles and are heterogeneous within each style; the fishing styles are based on an empirical classification that reflects the motivation as well as the practice of fishing (Boonstra & Hentati-Sundberg, 2016). The three fishing styles or types of fishers are the trawler fisher, the coastal fisher and the archipelago fisher, and the decision to fish is characterised by either bounded rational profit maximisation (trawler), bounded rational value maximisation (coastal), or bounded rational value satisficing (archipelago). The decision where to fish is knowledge-based for all styles, but the coastal and trawler fishers both use social information (descriptive norm).

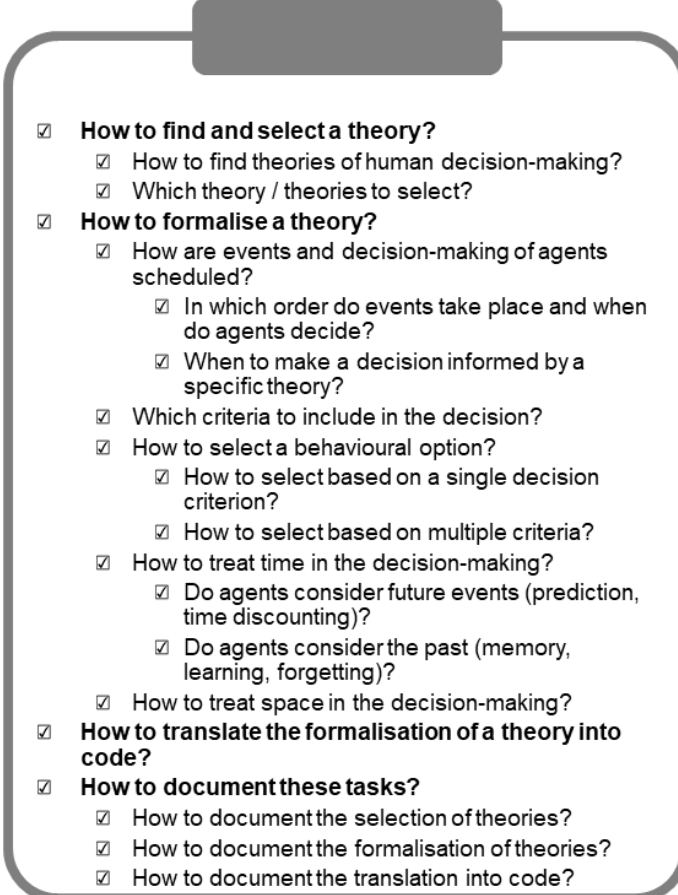
**Table 2:** Overview of the showcase models and respective implemented theories.

<b>Model</b>	<b>Content in a nutshell</b>	<b>Implemented (meta-) theories</b>	<b>Reasons for theory implementation</b>
Daisyworld (Janssen, 2016) Model code: <a href="https://www.comses.net/codebases/4958/releases/1.0.0/">https://www.comses.net/codebases/4958/releases/1.0.0/</a>	Stylised model using a generic resource	Bounded rationality / satisficing Descriptive norm Habitual behaviour / reinforcement learning Rational choice Theory of planned behaviour	Exploring the potential theory-dependent diversity in outcomes
FIBE (Wijermans et al., 2020) Model code: <a href="https://www.comses.net/codebases/4d9cfac5-7331-4a03-9d83-ab06cbcdc143/releases/1.0.0/">https://www.comses.net/codebases/4d9cfac5-7331-4a03-9d83-ab06cbcdc143/releases/1.0.0/</a>	Fishery model representing different fisher behaviours as described and classified in the (Swedish) Baltic Sea fishery	Bounded rationality / maximising Bounded rationality / satisficing Descriptive norm / social learning / imitation	Assessing the implications of motivational and behavioural diversity for the sustainability and management of the fishery
RAGE (Dressler et al., 2019) Model code: <a href="https://www.comses.net/codebases/5721/releases/1.0.6/">https://www.comses.net/codebases/5721/releases/1.0.6/</a>	Stylised model of resource use in a semi-arid rangeland system	Bounded rationality / satisficing Descriptive norm Rational choice	Representing agents on a spectrum between traditional, norm-abiding agents to more profit-oriented agents

The RAGE model (Dressler et al., 2019) simulates resource use in a stylised semi-arid rangeland system. The agents are pastoralist households raising livestock as their main asset, and all households have the same access rights to all pastures, i.e., a common pool resource. RAGE has been developed as a tool to understand how different human decision-making strategies influence long-term livestock and pasture conditions and whether a change in decision-making can drive the system into a degraded state or counteract such a development. In the RAGE model, three household behavioural types are included: a traditionalist type that values both livestock and pasture condition, a short-term profit maximiser type and a bounded-rational satisficer type.

### 3. Main tasks, experiences and recommendations

We focus on the four challenges of integrating theories of human decision-making into models as identified by Schlüter et al. (2017): the large number and diversity of theories, the focus of most theories on only a certain aspect of decision-making, their varying degree of formalisation, and the lack of specification of causal mechanisms. To address these challenges, we identified four main tasks specific to the inclusion of a particular theory of human decision-making in an ABM (Figure 1): 1) find and select a theory, 2) formalise it, 3) technically translate it into programming code, and 4) document its formalisation. The challenges and tasks do not form a 1-to-1 relationship: Task 1 of finding and selecting a theory clearly addresses the first challenge of the large number and diversity of theories. In addition, the degree of formalisation of a theory can be an important selection criterion. Furthermore, the aspect of decision-making that is covered by a theory is also relevant for its selection. Tasks 2 and 3 are closely linked to the challenging, varying degree of formalisation of theories. Also the need to formalise causal mechanisms is tackled in task 2, since modellers need to formalise how time is treated (what can lead to what?), which criteria are included in the decision-making et cetera. Task 4 on documenting the selection, formalisation and implementation results from the previous three tasks and, thus, addresses all challenges.

- 
- ☑ **How to find and select a theory?**
    - ☑ How to find theories of human decision-making?
    - ☑ Which theory / theories to select?
  - ☑ **How to formalise a theory?**
    - ☑ How are events and decision-making of agents scheduled?
      - ☑ In which order do events take place and when do agents decide?
      - ☑ When to make a decision informed by a specific theory?
    - ☑ Which criteria to include in the decision?
    - ☑ How to select a behavioural option?
      - ☑ How to select based on a single decision criterion?
      - ☑ How to select based on multiple criteria?
    - ☑ How to treat time in the decision-making?
      - ☑ Do agents consider future events (prediction, time discounting)?
      - ☑ Do agents consider the past (memory, learning, forgetting)?
    - ☑ How to treat space in the decision-making?
  - ☑ **How to translate the formalisation of a theory into code?**
  - ☑ **How to document these tasks?**
    - ☑ How to document the selection of theories?
    - ☑ How to document the formalisation of theories?
    - ☑ How to document the translation into code?

**Figure 1:** Main tasks and specific questions to be considered when including theories of human decision-making in ABMs.

Each of these four tasks can be approached with individual steps, which we formulated as questions and describe in the following sections. These steps belong to the model design and implementation phases of the modelling cycle (Schmolke et al., 2010). Including a particular theory of human decision-making can also influence later phases, and we provide an outlook on how theories can impact these phases in section 4.

### 3.1 How to find and select a theory?

For the task of finding and selecting one or more suitable theories for a given research question, we identified two questions: 1. How to find theories of human decision-making? 2. Which theory / theories to select? These questions address the first two challenges described by Schlüter et al. (2017), namely, the huge number of available theories of human decision-making and the different scopes of these theories. In the following sections, we share our experiences in how to find relevant theories and provide a list of potential sources (Table 3) and discuss factors that could be relevant for selecting a theory or several theories (summarised in Figure 2).

#### 3.1.1 How to find theories of human decision-making?

Finding theories of human decision-making for one's own modelling project is an enormous challenge, as different social science disciplines have developed a large variety of theories that could be applied to modelling the decisions of natural resource users (Schlüter et al., 2017). Modellers who already have knowledge about existing theories and the terminology used in psychology and economics could of course search for theories specifically addressing the specific modelling context. Modellers who are not familiar with the discipline-specific terminology are hindered to search for theories with appropriate keywords (for instance to find theories dealing with copying behaviour of others by using keywords such as “imitation” and “social norm”).

Thus far, these theories have not been collected into a single compendium or review; however, there are several existing compilations and other sources that may serve as a starting point (Table 3). Among them are social science encyclopaedias (Table 3) that contain theories on human decision-making, but also many other theories. Thus, modellers who are rather unfamiliar with the subject might find it hard to identify those theories that could be relevant for them. In addition to encyclopaedias, literature reviews, Wikipedia, textbooks and the like can provide an entry point; these can be related to environmental behaviour (e.g. Moore & Boldero, 2017; Table 3) but also about changing health-related behaviour (e.g. Garnett et al., 2018) can be relevant. Here, we focus on psychological and economic theories addressing individual decision-making. Natural resource use often involves common pool resources and collective decision-making (Scarlett et al., 2013). Several sources provided in Table 3 can also serve as starting points for investigating such theories (e.g. Nyblade et al., 2019; Turner et al., 2017; Scarlett et al. 2013).

Other ways of finding relevant theories include cooperating with experts on social science theories, attending context-related conferences, etc. While such contacts are very helpful for finding theories as well as selecting and formalising them, one must first find social science colleagues for such an interdisciplinary collaboration and build a common language. These options are good starting points for modellers seeking potentially relevant theories. However, what is needed is a FAIR (findable, accessible, interoperable, reusable) compilation of decision-making theories with case studies that are domain-relevant (e.g. for natural resource use).

In our joint modelling exercise, all modellers already had experience modelling human decision-making, so we knew where to begin looking for theories. Previously, we also conducted a literature survey of reviews on natural resource use ABMs to inform our framework for describing such theories for modelling purposes (Schlüter et al., 2017). Finally, the decision-making models for the different fishing styles in the FIBE model, i.e., the selection of the set of theories to combine to represent one style, were co-developed with creators of the empirically grounded generic concept of fishing styles reflecting another lens to classify fisher behaviour.

**Table 3:** Potential sources for finding theories of human decision-making.

Type of source	Main sources	Pros and cons
Compiled theories for modelling	Balke & Gilbert (2014) Brenner (2006) Jager (2017) Meyfroidt (2013)* Nyblade et al. (2019)	+ Focus is already on usability in models. - Example formalisations are not always available.
Summaries of theories used in ABMs	Alonso-Betanzos et al. (2017) An (2012) Cooke et al. (2009) Groeneveld et al. (2017)* Huber et al. (2018)* Klabunde & Willekens (2016)	+ Example formalisation already available. - Theories that could be useful but have never been used in a model are not included.
Compiled theories for policy makers	Darnton (2008) Dessart et al. (2019)* Scarlett et al. (2013) Van Vugt et al. (2014) World Bank (2015)*	+ Extensive overview. - Translation to modelling world needed.
Reviews of theories	Armitage & Conner (2000) Davis et al. (2015) Engler et al. (2019) Frederiks et al. (2015) Garnett et al. (2018) Moore & Boldero (2017) Steg & Vlek (2009)*	+ Extensive overview. + Highlighting most frequently used theories and / or quality criteria. - Translation to modelling world needed.
Introductory textbooks for behavioral economics, social and environmental psychology, decision theory	Camerer (2003) Coats (2001) Hewstone et al. (2015) Peterson (2017)* Steg et al. (2012)	+ Extensive overview. - Translation to modelling world needed.
Encyclopaedias of social sciences	Darity Jr. (2008) Ritzer (2012) Turner et al. (2017)	+ Extensive overview. - Hard to identify relevant theories - Translation to modelling world needed.
Informal methods	Cooperating with experts on social science theories Attending relevant conferences	+ Direct exchange about pros and cons of a specific theory, etc. possible. - Not always easy to find colleagues able to collaborate on a modelling project; need to find a common language.

Note: \* indicates key sources.

### 3.1.2 Which theory / theories to select?

Selecting theories is clearly context-dependent (also see Brenner, 2006, on selecting theories of learning, and Risjord, 2018, for nursing). According to Edmonds (2012), context not only relates to the specific social-ecological situation one wants to model but also to cognitive and social processes with which the individuals are involved as well as the linguistic framing of communication. Furthermore, some theories can instead be characterised as broader meta-theories with many different aspects that must potentially be considered. A well-known example is bounded rationality, which covers many concepts that can potentially be included in a model (e.g., the amount of information considered by the agents, processing capacity, errors in perception, Gigerenzer and Goldstein, 1996; Kahneman, 2003). While the modelling context influences theory selection, modellers will hardly find exactly one specific theory to match their problem. In the following, we discuss seven factors that may influence the selection of a theory (Figure 2) and, in doing so, aim to provide guidance for modellers new to the field.



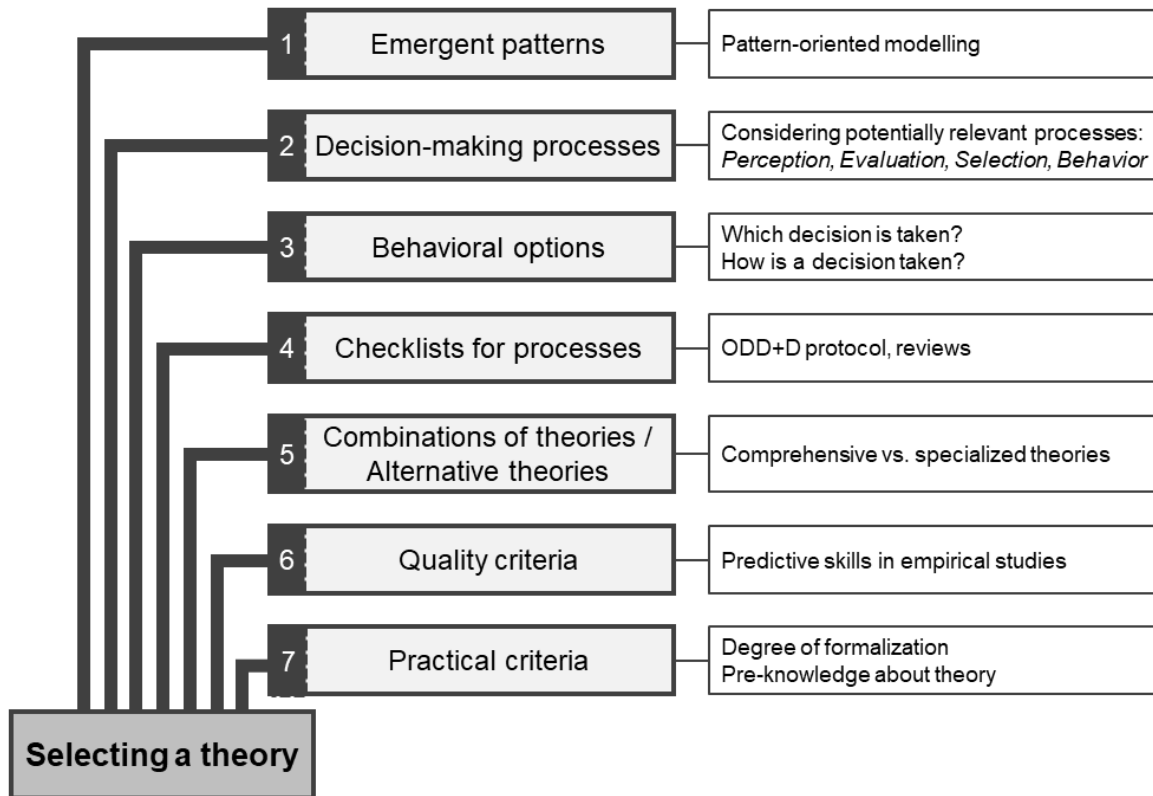


Figure 2: Relevant factors for selecting theories (left) with details that can be considered (right).

**Factor 1.** The emergent pattern(s) to be modelled (pattern-oriented modelling, Grimm et al., 2005) can favour a family of theories. For example, modelling social diffusion as an emergent pattern suggests theories with a focus on social interaction that address informative and normative social influence (see review in Cialdini & Goldstein, 2004) and theories of social learning and imitation processes (e.g., Bandura, 1977).

**Factor 2.** The processes of decision-making in a given model context can determine which theories are relevant. Building upon the MoHuB framework for mapping theories of human decision-making (Schlüter et al., 2017), the processes of perception, evaluation, selection and behaviour are potentially relevant for choosing one or several theories. In some contexts, perception, i.e., the process of actually sensing the environment, is highly important, such as when modelling intangible environmental risks (see also Milner-Gulland, 2012). Thus, the envisioned ecological dynamics in the model could play a role in choosing a theory, especially since we did not find many theories on human decision-making that consider the biophysical environment. The selection process can involve several types, including imitation, optimisation or satisficing (Schlüter et al., 2017) that can also hint at a specific theory; for instance, imitating others can be explained by the theory on imitation, persuasion, normative influence or social learning.

**Factor 3.** The behaviour and the available behavioural options to be included in the model may hint at certain theories. For example, modelling everyday decisions advocates theories on habits and how to overcome them or the use of fast and frugal heuristics (Gigerenzer & Goldstein, 1996). The latter suggests that decision problems are solved with simple rules focusing on only one aspect of the decision problem. Whether memory and learning must be considered in the decision process further helps to identify the appropriate theory; decision-making relying on memorised experience may be better described according to the approach of case-based reasoning (Aamodt & Plaza, 1994), which states that decision problems are solved based on experience with similar problems. This not only relates to how a decision is taken, but which one(s). For instance, behaviour aimed at reducing health risks suggests the Protection Motivation Theory (Maddux & Rogers, 1983, implemented, e.g., in Abdulkareem et al., 2018). This theory is actually a good example of a theory originally developed for a specific context (here, health risks) which was later applied in other fields (for instance, flood risk, Bubeck et al., 2012 or sustainable agriculture, Bopp et al., 2019). There are frameworks that help to find robust decision models under

uncertainty that are selected from candidate models describing different behavioural options (Lempert et al., 2006).

*Factor 4.* The ODD+D protocol (Müller et al., 2013) for describing human decision-making in ABMs can be used as a checklist to identify additional elements (such as uncertainty and risks, adaptation or learning, scheduling of updates, see also section 3.2.1.1) that are important to the purpose and context of the model. For instance, Prospect Theory (Kahneman & Tversky, 1979) is only applicable to contexts in which risks play a role. This is because Prospect Theory concerns rational decisions being biased because the context pushes risk aversion towards avoiding losses over chancing a gain. Additionally, categorising the factors that influence the evaluation (e.g., economics, social influence, social impact, environment, Groeneveld et al., 2017) may aid in identifying an appropriate theory.

*Factor 5.* Selecting several theories for a single model can also be an option, and this can be achieved by combining theories or treating them as alternatives. Combining them can be necessary in two situations: First, a single theory might not cover all aspects of the decision-making process that are relevant in the given model context (e.g. the theory focusses on perception of the natural resource, but is not explicit about the evaluation of the information). Second, the model context might require combining several decision-making processes, so that an integrated model is needed (Jager, 2017). For example, the Consumat approach addressed habits, imitation, social comparison and optimisation processes in a social-ecological model dealing with a renewable (fish) and non-renewable (mine) resources – depending on the satisfaction and uncertainty of an agent it will select one of the four possible decision processes in a given time step (Jager et al., 2000). Of course, no one theory can explain all human behaviours; some theories are broader in the sense that they combine different aspects, such as attitudes and social norms (Theory of Planned Behaviour), while others only focus on one component (e.g., descriptive norm). Thus, whether combining theories is necessary to explain the entire phenomenon depends on the model context and the explanatory power of the theories being considered. For example, the Theory of Planned Behaviour is quite comprehensive, but it still does not explicitly include the perception of the environment (Schlüter et al., 2017). Therefore, if the perception of the natural resource is important in a model, an additional theory might be required. In such cases, the selected theories should be philosophically compatible, for instance, not contradicting each other in basic assumptions. One example of such a contradiction would be combining a rational actor and habits for a single decision-making process, since they differ, for instance, with respect to awareness of the actual decision-making and processing of information.

Alternatively, one can include several theories as alternatives. First, this decision would depend on the purpose of the model and would be needed if the purpose is to understand the implications of alternative theories or behavioural types (Jager, 2000). Second, the model context could hint at using alternative theories if no available theories were clearly preferred (e.g., Janssen & Baggio, 2017), which forces modellers to be explicit about how decision-making is implemented (e.g. Grimm et al., 2005). Also, it could be relevant to check whether the theories under consideration are well-established and empirically validated (Edmonds, 2017); but again, the model purpose might exactly be to test and further develop a new theory.

*Factor 6.* Quality criteria pertaining to the theories could play a role, for instance how often they have been used in empirical studies and, based on these studies, their predictive skill. Such information is provided in quantitative theory reviews (e.g. Armitage & Conner, 2000; Garnett et al., 2018).

*Factor 7.* Practical criteria for selecting a theory must be considered such as the degree of formalisation of a given theory because more formalised theories leave fewer ambiguities for the modeller when formalising human decision-making (section 3.2). The availability of a formalisation of the theory in an existing model is also practical, and hence a model library (e.g. comses.net) can be consulted for successful implementations of the theory. Another practical criterion is the skills of the modeller as it is much easier to implement a theory with which one is familiar than a new complicated theory and its foundations, which may slow the diffusion of new theories.

For our showcase models, several theories of human decision-making were selected for various reasons. For the Daisyworld model, we deliberately chose an abstract model to implement five theories as an exercise to explore the potential diversity of outcomes when modelling natural resource use (see Table 2) using alternative theories of behaviour. For the FIBE model, several theories were implemented but as combinations. The purpose of combining several theories in the FIBE model was to describe the behaviour of three types of fishers observed

in the real world, and no single theory fulfilled that purpose. Thus, for each fisher type, at least two theories were combined to cover the decision of the agent on whether to go fishing and where to go. The subsistence “archipelago” fisher, for instance, follows a satisficing selection process, while the business-oriented trawler fisher maximises while considering the observed behaviour of others (descriptive norm and social imitation). They all represent a bounded rational model, but the types differ in their goals and selection procedures. The RAGE model includes different alternative theories to represent societal change, from traditional norm-abiding agents (descriptive norm) to more profit-oriented agents (rational actor) on the one hand to rather conservative, income-diversifying agents (bounded rational satisficer) on the other.

Furthermore, some theories were explicitly excluded. For instance, Prospect Theory was not included in any of the models as the modelling contexts did not suggest that risk was critically related to the choices being made. For the RAGE and FIBE models, the Theory of Planned Behaviour was not included as herding and fishing decisions are repeated and may often be habitual, so explicitly modelling intentions was considered inappropriate. For the FIBE model, the rational actor was not included, as the bounded rational form allowed for reflecting that the reality that fishers are constantly dealing with incomplete knowledge about their resource when deciding whether and where to fish.

In summary, in our experience many theories are available from which to choose; however reviews have shown that modellers tend to rely on very few theories, if they use a theory at all (Groeneveld et al., 2017; Huber et al., 2018). The choice of a theory can be guided by seven aspects: 1) What outcomes or patterns are of interest? What are the characteristics of the situation in which agents decide? 2) What decision-making processes (perception, selection, etc.) are important in the given decision situation? 3) What are the behavioural options of the agents? 4) Which decision-making elements are relevant to the purpose and context of the model, e.g., according to the ODD+D? 5) Should several theories be combined? 6) Which quality criteria should the theory meet? 7) Which practical considerations need to be taken into account, such as the availability of existing implementations? In the end, this choice is always subjective, so it is important to communicate which theories were selected and why in the model documentation and to assess the implications of these choices.

### 3.2 How to formalise a theory?

The task of formalising one or more theories for implementation in a simulation model refers to two key challenges highlighted by Schlüter et al. (2017): the varying degree of formalisation of different theories and their reliance on correlations rather than causalities. In total, we identified five overarching questions:

1. How are events and decision-making of agents scheduled?
2. Which criteria to include in the decision?
3. How to select a behavioural option?
4. How to treat time in the decision-making?
5. How to treat space in the decision-making?

The questions start with the scheduling (question 1) and then dive into the decision-making process of the agents. Questions 2 and 3 are related to elements of the MoHuB framework, while questions 4 and 5 relate to the temporal and spatial aspects identified in the ODD+D protocol. For several of these questions, we further developed sub-questions that could help in formalising theories. While we reflect on the choices modellers face and discuss our own approaches, we do not provide off-the-shelf solutions for formalising a theory.

One example for a theoretical construct that is hard to formalise in a simulation is the descriptive norm. In their original papers on a Focus Theory of Normative Conduct, Cialdini et al. (1990; 1991) explain the concept of a descriptive norm as the observed behaviour that most of the others perform versus the injunctive norm, describing one’s expectations of what is approved or disapproved by most others. The authors then use field experiments to employ interventions (e.g., anti-littering) to implement either one or both norms. Then, they analyse statistical models, such as chi-squared tests, to show the effect of descriptive and injunctive norms on observed behaviour (e.g., littering). The theory does not specify several aspects of the decision-making, so the modeller must make assumptions when formalising this theory. For instance, the conditions in which this imitating behaviour (i.e., following the littering behaviour of most of the others) will occur are not specified nor is the perception of the results of the behaviour (e.g., litter present in the area) or the identity of the others. Rangoni and Jager (2017) built upon the littering example to implement the Goal Framing Theory (Lindenberg & Steg, 2007), in which complying with norms is one out of three motives a person will follow. Furthermore, descriptive norm can be seen as a special case of imitation, since the behaviour of a majority of (similar) others

is imitated. Other variants of imitation are, for instance, imitating successful others (which might well be a minority of all agents in a simulation). Gotts and Polhill (2009) showed in simulation experiments that the conditions under which imitation occurs as well as the type of imitation have an effect on the success of farmer agents.

All our showcase models include the descriptive norm theory but in different ways (see also the pseudo-code in Table 4). In the Daisyworld model, the behaviour to be imitated is the choice of white or black daisies, and agents only use descriptive norm for their decision if they have a certain minimum level of utility; otherwise they do not imitate others. The Daisyworld agents do not consider all other agents in the world but derive information about the choices made by others in a neighbourhood of a certain radius that can vary (perception). In the FIBE model, when using the descriptive norm as social influence, the trawler and coastal fisher agents will, with a very high probability, go to the area where most other fishers go if other fishers are present. In only 20% of the time they will go to an area that they themselves remember as a good fishing ground in the past. In the RAGE model, the descriptive norm is used to determine whether agents will follow the pasture resting rule, and the implementation of descriptive norms follows the stylised model of Muldoon et al. (2014) and introduces an individual preference for following the resting rule and the factor of social susceptibility to following the

**Table 4:** Pseudo-codes for the implementation of descriptive norms in the showcase models.

Model	Pseudo-code
Daisyworld	<pre> if (utility &gt; minimum threshold) // agent is satisfied {     dominant-choice = colour of daisies consumed by                     majority of other agents in radius r      if (colour of daisy on patch == dominant-choice)         { consume daisy on patch }     else if (daisy on neighbouring patch == dominant-choice)         { consume daisy on neighbouring patch }         else { move to neighbouring unoccupied patch }     } else // agent is not satisfied {     if (daisy on patch available)         { consume daisy on patch }     else     {         if (daisy on neighbouring patch available)             { consume daisy on neighbouring patch }         }     } </pre>
FIBE	<pre> if (out == false &amp; other agents out == true) {     in 80% of cases: move to spot where most other agents fish     in 20% of cases: select spot based on own memory } else { select spot based on own memory } </pre>
RAGE	<pre> s // characteristic of each agent: social susceptibility to other agent's behaviour s ∈ [0, 1] q // characteristic of each agent: intrinsic preference for pasture resting s ∈ [0, 1]  descriptive-norm = proportion of all other agents following the resting rule Probability P of following the resting rule = s * descriptive-norm + (1 - s) * q  if (random number ∈ unif(0, 1) &lt; P ) { Select new pasture only from sufficiently rested pastures } else { Select from all pastures } </pre>

behaviour of others or staying with one's own tendency. The stronger the social susceptibility of an agent, the more likely it will follow the behaviour of the other agents instead of its own preference (see pseudo code in Table 4). When evaluating the behaviour of others, all other agents are considered. We will use these examples in the following sections whenever possible.

### 3.2.1 How are events and decision-making of agents scheduled?

#### 3.2.1.1 *In which order do events take place, and when do agents decide?*

In ABMs, more than one agent often acts in a single time step, which raises the question of what entity does what and in what order? In principle, multi-agent models differentiate between synchronous and asynchronous updates (Caron-Lormier et al., 2008) which can have strong implications for the model outcomes. For instance, two rational actors may find the very same solution to their optimisation problem, but only one can implement it (e.g. because a resource is being consumed by the other because it went first). Thus, it is up to the modeller to resolve the conflict. In the case of synchronous updating, this conflict has to be resolved explicitly; for example, actors could differ in power or auction-like processes to resolve the conflict. In asynchronous updating, the conflict is resolved implicitly since agents act one after the other, which results in path dependency. A special case of asynchronous updating is updating on the perceived need (Page, 1997). The method of updating is usually not informed by the decision theories that we have introduced thus far, since these theories are based on single individuals making decisions. Still, modellers need to make a choice on the updating and be aware of potential consequences. When synchronous updating is used, explicit processes to schedule agents' decision-making might also inform the theory selection process (section 3.1.2, factor 4).

Often, the order in which agents act is randomised at each time step to minimise the bias of asynchronous updating, but randomised asynchronous updating is a strong assumption of perfect equality between agents which does not fit with all model contexts. Nevertheless, in two of our showcase models, we implemented this update mode for the sake of simplicity (Daisyworld) and due to the lack of information on the hierarchy within the modelled population (RAGE). In FIBE, the fishers that are already at sea decide where to go first, and the others decide afterwards; the actual fishing is executed in a random order.

#### 3.2.1.2 *When to make a decision informed by a specific theory?*

Agents can make a decision at every time step, or their decision-making processes may be triggered, for instance, by specific events. Furthermore, a decision may not be taken in one time step but over several steps in the case of sequential decision-making (as formalised, e.g., in Ahn, 2010). Similarly, the execution of a behaviour that is informed by a specific theory might be related to specific events.

In our showcase models, agents make their decisions in each time step, including the decision not to act (e.g., to not go fishing in the FIBE model), but when an agent will behave according to a specific theory was implemented differently in our models. The Daisyworld model includes a threshold for imitation in the formalisation of descriptive norm; agents only imitate others if they are satisfied, i.e. their utility is above the threshold. Additionally, the FIBE model does not prescribe imitating other agents in all circumstances; here, the descriptive norm comes into play for the trawler and coastal fishers when they decide where to fish. Fishers go to the spot where most others go in 80% of the time; otherwise, they use their experiential knowledge of good spots for fishing in the past. The influence of imitating others in the RAGE model is not controlled via thresholds or random events but rather through the willingness of the individual to follow social influences. In all cases, we tried to prevent all agents from imitating each other, which may lead to artificial results.

### 3.2.3 Which criteria to include in the decision?

Many different criteria can be considered when modelling the decision of an agent. Bartkowski & Bartke (2018), for instance, provide an overview on characteristics influencing European farmers' decision-making (see also Dessart et al., 2019). The MoHuB framework groups the agent state into needs / goals, knowledge (including the social and ecological environment), assets and values, and some theories clearly specify which of those are accounted for in a decision. For example, in the descriptive norm theory, knowledge of the social environment is an important aspect of the state of an agent, particularly with respect to the dominant behaviour of other agents. However, the exact set of peers that shall be imitated is not specified, so this aspect was formalised

differently in our showcase models. The Daisyworld agents do not consider all other agents in the world but derive information about the choices made by others in a neighbourhood of a certain radius. In the FIBE model, the trawler and coastal fisher agents know about the location of all other fishers and fish where most others do. In the RAGE model, the behaviour of all other agents is considered during evaluation. These examples highlight the variety of options available when modelling social influence (for innovation diffusion: Kiesling et al. 2012).

In other cases, the theory will state how information is evaluated or assessed in general rather than specifically stating what will be evaluated and considered in a decision. For example, the rational actor approach implies that agents evaluate behavioural options with respect to their expected utility, but the theory does not specify which criteria contribute to utility and are thus relevant for a given decision as they depend on the model context. In our showcase models, for instance, this context-dependency led the utility function of the rational actor to only include the state / amount of the natural resource (i.e., pasture biomass) used in one model (RAGE), but in another model, the utility function includes additional factors such as leisure time (Daisyworld).

Another example from the showcase models is the formalisation of satisficing that differs greatly with respect to the included criteria. In the Daisyworld model, satisficing agents have a threshold for their utility; agents will only act if their current utility is below that threshold. If their utility is above the threshold, the agent stays and does not perform the behaviour. In the FIBE model, the archipelago fisher agents use a satisficing selection process when deciding to fish, i.e., they reflect on whether they have fished enough (are satisfied). More specifically, for this threshold, they consider the costs of a five-working-day week and determine whether these costs have already been covered by the profit from the current week. If their financial capital is below zero or if their experience indicates that there is not enough fish, they will either go or not go fishing, respectively, independent of their catch satisfaction. Satisficing in the RAGE model depends on how many animals herders will stock; the satisficer agent type only stocks up to a predefined number that satisfies his needs.

In summary, many theories will not prescribe the exact criteria and their values to be included when modelling the decision of an agent. Even if a theory states the criteria that should be included, their exact formalisation will likely remain open. The model context might influence the choices that must be made. This highlights the importance of combining theory, empirical case knowledge and modelling choices in defining the criteria. As an example, the choice of observing the behaviour of others will depend on the specifics of the case. In the FIBE model, it is a reasonable assumption that fishers can perceive all other fishers, as the fishers' electronic systems share information on the destination, current location and speed of all vessels. Hence, in modelling a specific case one should decide on how a theory would work out in a given case, and how this can be captured in computational rules. In any case, model communication and transparency would be greatly benefitted by modellers clearly documenting and explaining their choices.

### 3.2.4 How to select a behavioural option?

#### 3.2.4.1 *How to select a behavioural option based on a single decision criterion?*

If the selection of a behavioural option is based on a single decision criterion or an aggregate variable, such as utility (see section 3.2.4.2 for how to compute such aggregate variables), the final selection can be performed by optimising or applying a threshold, and this is often done by selecting the option that scores best. For instance, the rational actor approach is often interpreted as utility maximisation, meaning that agents choose the behavioural option with the highest utility for them, and this interpretation of utility and the underlying preference is called a vector model, meaning that the higher the value, the more the agent benefits. However, utilities can also be ideal point types, meaning that a specific quantity (e.g., salt in food) maximises utility, and both under- and overshooting this quantity reduces utility (see Jager, 2007, for more examples on vector versus ideal point utilities). If other rules such as thresholds are applied, evaluating behavioural options instead means creating a subset of behavioural options that fulfil the threshold(s). In those cases, additional rules must specify the behavioural option to be selecting from that subset.

In our showcase models, different alternatives were chosen to select a behavioural option. For instance, the rational actors in the Daisyworld and RAGE models choose the behavioural option with the highest utility, and the trawler and coastal fishers also follow a maximising strategy (e.g., by maximising profit and/or time at home). In many other cases in our showcase models, decision trees with thresholds (see the next section) were used to determine a subset of feasible behavioural options, from which one behavioural option is randomly chosen. For

example, satisficer agents in the RAGE model select the first pasture that allows them to sustain their satisficing level of livestock out of a predefined subset of all pastures.

### 3.2.4.2 How to select a behavioural option based on multiple decision criteria?

If more than one criterion has been identified as relevant to the decision-making process, formulas for aggregate variables or rules need to be established before a behavioural option can be selected. For some theories, how the relevant criteria are to be included in the decision is quite open, and it is up to the modeller to decide on the type of equation or rule. Other theories are more formalised but still require some assumptions.

For example, the rational actor approach assumes that the agent evaluates the utility of each behavioural option as an aggregate variable, but even for such a highly formalised theory, several versions of the utility function exist, the most common being additive multi-attribute utility functions (weighted sum of all criteria, Equation 1) or the Cobb-Douglas utility (weighted product, Equation 2). The difference in formalisation leads to two criteria being included into the utility function separately (Equation 1) or jointly (Equation 2). This implies that a positive, non-zero value for utility can only be achieved for Equation 2 if all criteria score greater than zero, while zero values for a single criterion can still lead to a positive evaluation in Equation 1.

$$\text{Equation 1: } \text{Utility}(i) = \alpha * x(i) + \beta * y(i)$$

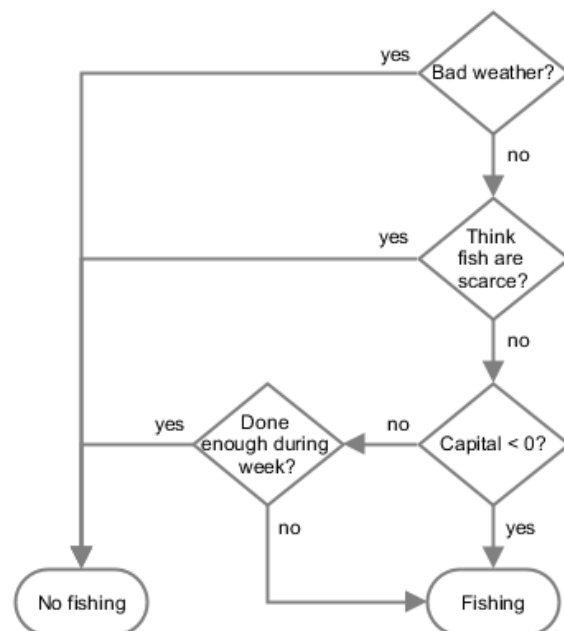
$$\text{Equation 2: } \text{Utility}(i) = x(i)^\alpha * y(i)^\beta$$

with  $x$  and  $y$  being criteria to be considered when deriving the utility of an option  $i$ ,  $\alpha$  and  $\beta$  preferences for criteria  $x$  and  $y$ , respectively. Often, it is required that  $\alpha + \beta = 1$ .

Although commonly used, combining multiple decision criteria into a utility function is just one of many options for a multi-criteria decision (e.g., Belton & Stewart, 2002). Generally, compensatory and non-compensatory rules can be distinguished, and these options differ in how the overall value of one behavioural option is affected by a low value in one of the determining criteria (Elrod et al., 2004). Non-compensatory decision rules — as their name suggests — do not allow for any compensation. For instance, if a threshold for a criterion is not met by one behavioural option, this option does not pass. Both types of rules can be combined in a decision tree to formalise a stepwise decision.

In our showcase models, many different options were chosen to treat multiple decision criteria. The rational actor in the Daisyworld model uses a multiplicative Cobb-Douglas utility function as this is a classical representation, but in many other parts of the showcase models decision trees were used to set up rules for linking multiple criteria. For example, the archipelago fisher in the FIBE model considers the fish stock, its capital, the weather and the amount of fishing done in the current week in the decision to go fishing (Figure 3).

To sum up, many different options are available for selecting a behavioural option based on single or multiple decision criteria. We believe it is important to be aware of the choice and the implications, for example, of compensatory versus non-compensatory approaches as summarised above. While this seems to be a little detail not worth discussing, it could in fact greatly influence outcomes.



**Figure 3:** Visualisation of a decision tree in the FIBE model: an archipelago fisher deciding whether to fish or not (modified after Wijermans et al. 2020).

### 3.2.5 How to treat time in the decision-making?

#### 3.2.5.1 Do agents consider future events (prediction, time discounting of future outcomes)?

Agents may predict future events (such as the future behaviour of others, future resource levels, climate conditions and market expectations), and how they assess future events or risks has fundamental consequences for their decisions; for example, how the discounting of future outcomes is modelled drastically changes the expected utility (O'Hare et al., 2009). When agents predict future events, computational restrictions also come into play. For example, it is practically impossible to formalise a *Homo economicus* with expected utility in an ABM, as decisions of other agents in future time steps cannot be completely considered.

None of our showcase models includes prediction of future events. Even in the formalisation of the rational actor, utilities relate to the current time step. However, it is possible to let agents make forecasts on expectations of, for example, future harvests and market-prices on historical data, and agents may differ concerning the length of the historic data period they consider in predicting future events (Speelman, 2014).

#### 3.2.5.2. Do agents consider the past (memory, learning, forgetting)?

Agents may consider the past by acquiring a memory of past events and using this information in later time steps, which can even lead to learning, defined here as a change in the decision rules themselves over time (Müller et al., 2013). Furthermore, forgetting impacts the historical timespan being considered by the agent in decision-making.

Of our showcase models, Daisyworld and FIBE encompass some form of agent memory. In the Daisyworld model, agents engaging in habitual behaviour or reinforcement learning have a memory of the past success of their behavioural options, which is formalised as a propensity to choose a specific behavioural option. This propensity is equal for all options in the first time step and updated according to the actual utility acquired when choosing a specific option. Thus, the agents do not remember specific events but the best result derived from each chosen behavioural option. In the FIBE model, all agent types have a memory of good fishing spots, i.e., places where they encountered enough fish in the past. Fishers may vary in how many locations they can remember, which affects how the past catch influences expectations of fishing locations.

In summary, several issues must be addressed when formalising theories for ABMs as temporal aspects are often not exactly prescribed by a theory. Many theories are based on statistical relations, and the static correlational relations between concepts are not easily translated into dynamics that develop over time. Again, the model context might influence these choices, and clear documentation would be helpful. All temporal aspects are mentioned in the standardised protocol for describing human decision-making (Müller et al., 2013).

### 3.2.6 How to treat space in the decision-making?

Most models of social-ecological systems are spatially explicit and include, for instance, heterogeneity in environmental properties such as resource availability. Also, agents are often distributed in space and act in specific localities. When explicitly modelling human decision-making, modellers need to clarify the spatial extent and resolution of their model, which influence the institutional setting and, thus, the actors relevant to the system, for instance, individuals, a community, a municipality, et cetera. Also, the actual process of decision-making needs to be formalised with respect to space. This can be challenging, since many theories on human decision-making are not explicit about aspects related to space. When formalising theories for our showcase models, we encountered the following challenges.

First, the perception of agents needs to be formalised with respect to space: Do agents sense the whole environment or only a specific area? For a boundedly rational agent, imperfect information could mean that the agent does not perceive the whole environmental system represented in the ABM. Also, imperfect information could imply that agents only sense other agents in a specific area, since social networks can also represent Euclidian space. The latter could, for example, be relevant for descriptive norm and the social norm in the Theory of Planned Behavior. For both, the researcher needs to formalise who and where the (important) others are whose decisions or opinions are perceived by the agents. Here we also need to consider that normative influence may be more prevalent in local physical networks, whereas informative influences can also operate in a wider



network context using e.g. the internet. For expected utility, a restricted perception could imply that the optimisation only considers a subset of all options.

Second, the processes of evaluation and selection of behavioural options could depend on space, and the challenge is to identify such spatial features in case the theory is not explicit about them. Such spatial factors could be, for instance, travel costs (e.g., financially, time, energy needed). Spatial factors can also manifest themselves as constraints limiting the options of agents. The Theory of Planned Behavior encompasses perceived behavioural control, which could encompass such spatial factors.

Third, the behaviour could be spatially explicit: Movement is an obvious spatially explicit behaviour, but ABMs could also entail behaviour influencing space in other ways, for instance by altering cell characteristics, spill-over effects onto other cells, et cetera. Typically, the spatial manifestation of behaviour is not covered by theories on human decision-making. Thus, the challenge lies in identifying such spatially explicit behaviour and formalising its specific parameters if needed.

All our showcase models are spatially explicit, and space is influencing agents' decision-making to various degrees, from a strong emphasis on local effects in the Daisyworld model to global perception of environment and agents in the RAGE model. In the Daisyworld model, agents' perception of the environment and other agents is limited to their own cell and the neighbouring cells within a predefined radius. For evaluating options according to the Theory of Planned Behavior, distance is inversely related to behavioural control. For habitual behaviour or reinforcement learning, an empty neighbouring patch is added as an option to explore. The agents' behaviour, i.e. the harvesting of daisies, is local with global effects for the temperature. In the FIBE model, fishers' perception of the environment is limited to the fish stocks in their vicinity, depending on whether they have the technology (e.g. radar), they can see the fish in the patch they are standing on or also the patches connected to their current location. Fishers perceive where others fish, however depending on their style make use of that information. When evaluating their options, fishers also indirectly consider space through travel costs to the different regions. Indeed, the options fishers decide upon are explicitly spatial, since if they fish, they decide where to fish. Their behaviour has a local effect on the fish stock. In the RAGE model, agents perceive the state of all pastures (i.e. available biomass for grazing) and consider the behaviour of all households (i.e. whether they rest pastures) in the descriptive norm. Space does not influence the evaluation or selection of behavioural options. Agents' behaviour is local, since their livestock only feeds on the pasture where it is currently located, and neighbouring patches do not influence each other.

In summary, space enters agents' decision-making in many ways. Many choices taken during formalising a theory, for instance related to perception, are implicit but still can be very influential. Therefore, it is important to be clear about them and document the choices and the reasoning behind.

### 3.3 How to translate the formalisation of a theory into code?

The steps mentioned above could lead for instance to pseudo-code or graphical representations such as decision trees or flow charts. However, there is still another step needed, namely implementing the formalisation into actual programming code. This includes taking technical decisions such as integer versus float, data structures and the choice of the random generator (e.g. Poile & Safayeni, 2016). These decisions may influence simulation outcomes and severely affect reproducibility (Edmonds & Hales, 2003). This is a topic that all simulation models have in common, not only models of human decision-making.

The NetLogo modelling platform (Wilensky, 1999) was used to implement all our showcase models and run the simulation experiments. NetLogo provides a large set of predefined functions (e.g. 'ask', 'move', or 'one-of') that follow a set logic, which can potentially influence simulation outcomes, for instance by encouraging a random sequencing of agents' behaviour. A general purpose language (Java, C++, etc.) would have required the modellers themselves to specify all these logics.

### 3.4 How to document the selection, formalisation and translation?

The last three sections have shown that a) there are many different theories to choose from, and many ways exist to b) formalise any given theory and c) translate that into code. Thus, documentation of a theory incorporated into an ABM would ideally be very specific about these three tasks.

For ABMs, several types of documentation are available (Müller et al., 2014):

- Natural language descriptions
  - Prescriptive structure such as that given by the ODD+D (Müller et al., 2013) and ODD+2D (Laatabi et al., 2018) protocols, both based on the ODD (Grimm et al., 2020).
  - Non-prescriptive structure, i.e., a verbal description with no specific order
- Formal language descriptions
  - Ontologies that formally describe entities and their relationships
  - Source code
  - Pseudo-code combining natural language with programming syntax
  - Mathematical descriptions such as in the form of equations
- Graphics
  - Formal graphics that, for example, follow the Unified Modelling Language
  - Non-formal graphics such as flow charts

In the next sections, we discuss which types of documentation can be useful for documenting the different tasks and suggest specific documentation types for each task.

### 3.4.1 How to document the selection of theories?

To state which theory has been selected and the reasons for that, natural language descriptions are a good option. The structured model documentation protocol ODD+D for describing human decision-making in ABMs (Müller et al., 2013) explicitly asks for the theoretical background of the model. The ODD+2D protocol (Laatabi et al., 2018) expands this protocol with a part about the use of data in these models. Additionally, graphics can be used to justify the combination of several theories in one model, such as building on the MoHuB framework (Schlüter et al., 2017).

In our showcase models, we justified our choices for specific theories in the main bodies of the respective publications (Table 2). We provided reasons for selecting these theories, but no reasons for why others related were not chosen since the documentation would otherwise become too lengthy.

### 3.4.2 How to document the formalisation of theories?

As already explained above, documenting the formalisation itself also needs to report on the many little details (and choices) that modellers typically make on their own that are not prescribed by the theory, such as whether multiple decision criteria were multiplicatively or additively combined in a utility function. Furthermore, it would be very helpful to not only describe the final formalisation but the underlying reasons for certain choices as well, such as why a multiplicative or additive function was used. To describe the formalisation itself, source code, pseudo-code or mathematical descriptions are potential candidates. Ontologies would likely describe entities and their relationships but not, for instance, the way multiple decision criteria are used in the selection process.

In our showcase models, we rarely exchanged the source code itself when discussing the formalisation of theories – even though we had implemented all models in the same programming environment of NetLogo. Rather, we often communicated our theory formalisations using graphical representations such as decision trees (Figure 3) or graphs building on the MoHuB framework (Schlüter et al. 2017). We also discussed formalisation details using mathematical descriptions and pseudo-code (Table 4), and we used the ODD+D protocol to describe the models and the theoretical background as appendices to the scientific publications on the individual models.

### 3.4.3 How to document the translation into code?

Apart from documenting the formalisation, it would be even more helpful for other modellers to publish the source code itself, which could eventually lead to shared model components in a model library such as comses.net (Bell et al., 2015; Rollins et al., 2014). Ultimately, only the actual source code can provide all technical details such as data types and the like (Poile & Safayeni, 2016). All our showcase models are provided as open source software on comses.net. For any modelling project we recommend exploring the library of existing models in the comses.net for implementations that address issues that are relevant in the context of the project at hand. Often different components of existing models can be used, sometimes with some adjustments. This is

a good practice especially when models have been used in scientific publications, and the code has been reviewed externally (starred models).

#### 4. Theory-related issues in later stages of the modelling cycle

When incorporating theories of human decision-making, additional aspects must be addressed in later stages of the modelling cycle, but are beyond the scope of this paper. First, these models need to be parameterised and calibrated as the decision-making will likely account for agent-related parameters such as preferences and attitudes as well as other attributes. The very basic way of establishing these parameters determines the level of agent heterogeneity in the model, which can range from completely identical agents to different agent types, with varying properties even within a single type (Brown & Robinson 2006). Assumptions or data are used to determine these parameters, either for one generic agent or different agent types. Useful references are Smajgl and Barreteau (2017) who provide a framework for parameterising agents using different data sources, and others who demonstrated parameterisation based on empirical data analysed in a Bayesian framework (e.g., Pope & Gimblett, 2015) and inverse modelling using Bayesian inference (Hartig et al., 2011).

Second, the method of analysing model results could depend on the underlying theory of human decision-making. A model of deterministically, well-informed, optimising agents requires a different analytical treatment than one of agents with limited information that face uncertainty. Highly stochastic models need more demanding sampling strategies than deterministic models to determine the sensitivity of the model outcomes to the parameters (e.g., Thiele et al., 2014).

Third and related to the previous aspects, a sensitivity analysis that explicitly considers the assumptions made when incorporating a theory is demanding. All the choices and assumptions that a modeller has to make during the process of formalising one or several decision theories require tests of the robustness of the results to key assumptions.

#### 5. Conclusions

In this paper, we highlighted the possibilities, difficulties and choices one faces when incorporating one or several theories on human decision-making in a social-ecological ABM. Using three showcase models of natural resource use, we described the process of including a theory into an ABM by completing four tasks, namely, theory selection, formalisation, translation into code and documentation. For each task, we provided a systematic overview of steps we had to take for our models and highlighted the choices we made and pointed to useful resources to provide guidance for fellow modellers. We showed that the formalisation of a theory depends to some extent on the context in which the model is embedded. Because of the inevitable variety in formalisations and implementations, meaningful and unambiguous documentation is of the utmost importance for others to understand exactly how human decision-making has been formalised and implemented and why a specific approach was chosen. We confirmed that no single type of documentation can fulfil these purposes (see Müller et al., 2014), so we recommend a combination of a verbal explanation and at least pseudo-code and preferably actual source code. This verbal documentation ideally not only includes a description of the selection, formalisation and implementation as such, but also the reasons for choices made during the process.

The process of selecting, formalising, implementing and documenting theories of human decision-making in ABMs will be more transparent if more modellers share their experiences and the reasoning behind their choices with the scientific community. Those details can be documented within appendices using documentation protocols such as the ODD protocol (Grimm et al. 2020), including the ODD+D version of the protocol which explicitly highlight human decision-making (Müller et al., 2013). Besides improved model documentation, it is important that scholars share their model code, which currently happens in less than 20% of the papers using agent-based modelling (Janssen, 2017). By sharing the technical details and the motivations of the choices we hope that this will enable modellers to learn from one another and ultimately help develop shared practices for theory-informed modelling of human behaviour in social-ecological models. A broader compilation of existing theories of human decision-making and their implementations would be very helpful for modellers facing the task of selecting theories for their modelling context. In our work, we focussed on behavioural economics and social and environmental psychology. Future work on a broader compilation of theories should also consider

social science disciplines concerned with groups such as sociology or anthropology to also capture how individual decision-making gives rise to collective patterns. A critical benefit of providing documentation details of implementations and code sharing is the opportunity for reusing model components so that they can be better tested and results compared across contexts. It also facilitates comparison of the effects on outcomes of different decision theories or implementations thereof. Moreover this allows for a critical review of implementations, which further contributes to the development of a high quality repository of implementations, as in CoMSES.net.

Modelling social-ecological systems is often performed in interdisciplinary teams. These teams can be instrumental for making well-informed choices when implementing human decision-making theories in an ABM, particularly if the team involves social scientists familiar with the theories. However, our collective experience in diverse interdisciplinary projects indicates that more can be done (Davis et al., 2018). Involvement of diverse empirical social scientists in the conception and implementation of formal models, and exposure of modellers to the design and implementation of empirical social science studies, may increase mutual understanding of opportunities and limitations of modelling human decision-making. However, within psychology and economics, ABMs still receive rather limited attention (but see e.g. Smith and Conrey, 2007; Gräbner et al., 2017). Hence it is also a challenge to educate a new generation of social scientists in the relevance of this methodology in understanding social and socio-ecological dynamics (Jager et al., 2020). Such experiences could be gained by special training like summer schools, where colleagues from diverse disciplinary backgrounds are brought together to work on team projects. Increasing attempts in the social sciences to provide formalised representations of theories (West et al., 2019) will further help in this endeavour. A tighter collaboration between the modellers and social scientists is needed and could lead to a better understanding of human-environmental interactions.

## Acknowledgements

We gratefully acknowledge the financial support from the National Socio-Environmental Synthesis Center in Annapolis, MD, US (SESYNC); the Helmholtz Centre for Environmental Research (UFZ) in Leipzig, Germany; and the German Centre for Integrative Biodiversity Research (iDiv), Leipzig for hosting meetings of our working group. We also thank our colleagues Andres Baeza, Ryan McAllister, Kirill Orach and Nathan Rollins within the “Human Decisions & Ecosystem Services” project of SESYNC for inspiring discussions during our research. M.S. and N.W. acknowledge funding by the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP/2007-2013/ERC grant agreement no. 283950 SES-LINK), by the ERC under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 682472 — MUSES) and a core grant to the Stockholm Resilience Centre by Mistra. B.M. and G.D. were supported by the German Federal Ministry of Education and Research (BMBF—01LN1315A) within the Junior Research Group POLISES. We also acknowledge the valuable feedback from the students of the winter school in January 2017 on “How to model human decision-making in social-ecological agent-based models” in Tempe, Arizona, US and from participants of the workshop “Using theories on human decision-making in ABMs” at the Social Simulation Conference 2018 in Stockholm, Sweden.

Author contributions: All authors contributed to the design of the research. G.D., N.W. and M.A.J. are the lead authors of the showcase models; N.S. and J.G. led the writing of the paper; and all authors contributed to the processes of writing and discussion.

## References

- Aamodt, A., & Plaza, E. (1994). Case-Based Reasoning: Foundational Issues, Methodological Variations and System Approaches. *AI Communications*, 7(1), 39–59. <https://doi.org/10.3233/AIC-1994-7104>
- Abdulkareem, S. A., Augustijn, E. W., Mustafa, Y. T., & Filatova, T. (2018). Intelligent judgements over health risks in a spatial agent-based model. *International Journal of Health Geographics*, 17(8), 1–19. <https://doi.org/10.1186/s12942-018-0128-x>
- Ahn, H. (2010). Modeling and Analysis of Affective Influences on Human Experience, Prediction, Decision Making, and Behavior. Massachusetts Institute of Technology. <http://hdl.handle.net/1721.1/61929>
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)

- Ajzen, I. (2012). Martin Fishbein's legacy: The reasoned action approach. *Annals of the American Academy of Political and Social Science*, 640(1), 11–27. <https://doi.org/10.1177/0002716211423363>
- Alonso-Betanzos, A., Sánchez-Marroño, N., Fontenla-Romero, O., Polhill, J. G., Craig, T., Bajo, J., & Corchado, J. M. (2017). *Agent-Based Modeling of Sustainable Behaviors*. Springer. <https://doi.org/10.1007/978-3-319-46331-5>
- An, L. (2012). Modeling human decisions in coupled human and natural systems: Review of agent-based models. *Ecological Modelling*, 229, 25–36. <https://doi.org/10.1016/j.ecolmodel.2011.07.010>
- Anderson, C. (2008). The End of Theory, Will the Data Deluge Makes the Scientific Method Obsolete? [online] Available at: <https://www.wired.com/2008/06/pb-theory/>
- Armitage, C. J., & Conner, M. (2000). Social cognition models and health behaviour: a structured review. *Psychology and Health*, 15, 173–189.
- Balke, T., & Gilbert, N. (2014). How Do Agents Make Decisions? A Survey Introduction: Purpose & Goals Dimensions of Comparison. *Journal of Artificial Societies and Social Simulation*, 17(4), 13. <https://doi.org/10.18564/jasss.2687>
- Bandura, A. (1977). *Social learning theory*. Prentice-Hall.
- Bartkowski, B., & Bartke, S. (2018). Leverage points for governing agricultural soils: A review of empirical studies of European farmers' decision-making. *Sustainability*, 10, 3179. <https://doi.org/10.3390/su10093179>
- Bell, A. R., Robinson, D. T., Malik, A., & Dewal, S. (2015). Modular ABM development for improved dissemination and training. *Environmental Modelling and Software*, 73, 189–200. <https://doi.org/10.1016/j.envsoft.2015.07.016>
- Belton, V., & Stewart, T. (2002). *Multiple Criteria Decision Analysis*. Springer.
- Bopp, C., Engler, A., Poortvliet, P. M., & Jara-rojas, R. (2019). The role of farmers' intrinsic motivation in the effectiveness of policy incentives to promote sustainable agricultural practices. *Journal of Environmental Management*, 244, 320–327. <https://doi.org/10.1016/j.jenvman.2019.04.107>
- Boonstra, W. J., & Hentati-Sundberg, J. (2016). Classifying fishers' behaviour. An invitation to fishing styles. *Fish and Fisheries*, 17(1), 78–100. <http://doi.org/10.1111/faf.12092>
- Brenner, T. (2006). Agent learning representation: advice on modeling economic learning. *Handbook of Computational Economics*, 2, 895–947. [https://doi.org/10.1016/S1574-0021\(05\)02018-6](https://doi.org/10.1016/S1574-0021(05)02018-6)
- Brown, D. G., & Robinson, D. T. (2006). Effects of heterogeneity in residential preferences on an agent-based model of urban sprawl. *Ecology and Society*, 11(1), 46. <http://www.ecologyandsociety.org/vol11/iss1/art46/>
- Bubeck, P., Botzen, W. J. W., & Aerts, J. C. J. H. (2012). A Review of Risk Perceptions and Other Factors that Influence Flood Mitigation Behavior. *Risk Analysis*, 32(9), 1481–1495. <https://doi.org/10.1111/j.1539-6924.2011.01783.x>
- Camerer, C. F., Loewenstein, G., & Rabin, M. (Eds.). (2003). *Advances in Behavioral Economics*. Princeton University Press.
- Caron-Lormier, G., Humphry, R. W., Bohan, D. A., Hawes, C., & Thorbek, P. (2008). Asynchronous and synchronous updating in individual-based models. *Ecological Modelling*, 212(3–4), 522–527. <https://doi.org/10.1016/j.ecolmodel.2007.10.049>
- Cialdini, R. B., & Goldstein, N. J. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, 55(1974), 591–621. <https://doi.org/10.1146/annurev.psych.55.090902.142015>
- Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A Focus Theory of Normative Conduct: A Theoretical Refinement and Reevaluation of the Role of Norms in Human Behavior. *Advances in Experimental Social Psychology*, 24, 201–234. [https://doi.org/10.1016/S0065-2601\(08\)60330-5](https://doi.org/10.1016/S0065-2601(08)60330-5)
- Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: Recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, 58(6), 1015–1026. <https://doi.org/10.1037/0022-3514.58.6.1015>
- Coats, E. J. (2001). *Classic and Contemporary Readings in Social Psychology*. Pearson.
- Cooke, I. R., Queenborough, S. A., Mattison, E. H. A., Bailey, A. P., Sandars, D. L., Graves, A. R., ... Sutherland, W. J. (2009). Integrating socio-economics and ecology: A taxonomy of quantitative methods and a review of their use in agro-ecology. *Journal of Applied Ecology*, 46(2), 269–277. <https://doi.org/10.1111/j.1365-2664.2009.01615.x>
- Crooks, A., Castle, C., & Batty, M. (2008). Key challenges in agent-based modelling for geo-spatial simulation. *Computers, Environment and Urban Systems*, 32(6), 417–430. <https://doi.org/10.1016/j.compenvurbsys.2008.09.004>
- Darity, W.A. (ed., 2008). *International encyclopedia of the social sciences*. Macmillan.
- Darnton, A. (2008). *Practical guide: An Overview of behaviour change models and their uses*. London. [https://doi.org/https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/498065/Behaviour\\_change\\_reference\\_report\\_tcm6-9697.pdf](https://doi.org/https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/498065/Behaviour_change_reference_report_tcm6-9697.pdf)
- Davis, P. K., O'Mahony, A., Gulden, T. R., Osoba, O. A., & Sieck, K. (2018). Priority Challenges for Social and Behavioral Research and Its Modeling. Retrieved from [https://www.rand.org/pubs/research\\_reports/RR2208.html](https://www.rand.org/pubs/research_reports/RR2208.html)
- Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing theory through simulation methods. *The Academy of Management Review*, 32(2), 480–499. <https://doi.org/10.5465/AMR.2007.24351453>
- Davis, R., Campbell, R., Hildon, Z., Hobbs, L., & Michie, S. (2015). Theories of behaviour and behaviour change across the social and behavioural sciences: a scoping review. *Health Psychology Review*, 9(4), 323–344. <https://doi.org/10.1080/17437199.2014.941722>
- Dessart, F. J., Barreiro-Hurlé, J., & Bavel, R. Van. (2019). Behavioural factors affecting the adoption of sustainable farming practices: a policy-oriented review. *European Review Of Agricultural Economics*, 46(3), 417–471. <https://doi.org/10.1093/erae/jbz019>

- Dressler, G., Groeneveld, J., Buchmann, C. M., Guo, C., Hase, N., Thober, J., Frank, K. & Müller, B. (2019). Implications of behavioral change for the resilience of pastoral systems—Lessons from an agent-based model. *Ecological Complexity*, 40(B), 100710. <https://doi.org/10.1016/j.ecocom.2018.06.002>
- Edmonds, B. (2012). Context in social simulation: why it can't be wished away. *Computational and Mathematical Organization Theory*, 18(1), 5–21. <https://doi.org/doi:10.1007/s10588-011-9100-z>
- Edmonds, B. (2017). The Post-Truth Drift in Social Simulation. In *Social Simulation 2017*. Dublin. Retrieved from [http://cfpm.org/file\\_download/179/SSC\\_Edmonds\\_SSC2017-309.pdf](http://cfpm.org/file_download/179/SSC_Edmonds_SSC2017-309.pdf)
- Edmonds, B., & Hales, D. (2003). Replication, replication and replication: Some hard lessons from model alignment. *Journal of Artificial Societies & Social Simulation*, 6(4), 11. Retrieved from <http://jasss.soc.surrey.ac.uk/6/4/11.html>
- Elrod, T., Johnson, R. D., & White, J. (2004). A new integrated model of noncompensatory and compensatory decision strategies. *Organizational Behavior and Human Decision Processes*, 95(1), 1–19. <https://doi.org/10.1016/j.obhdp.2004.06.002>
- Elsawah, S., Filatova, T., Jakeman, A. J., Kettner, A. J., Zellner, M. L., Ioannis, N., Hamilton, S. H., Axtell, R. L., Brown, D. G., Gilligan, J. M., Janssen, M. A., Robinson, D. T., Rozenberg, J., Ullah, I. I. T., & Lade, S. J. (2020). Eight grand challenges in socio-environmental systems modeling. *Socio-Environmental Systems Modelling*, 2, 16226. <https://doi.org/10.18174/sesmo.2020a16226>
- Engler, J. O., Abson, D. J., & von Wehrden, H. (2019). Navigating cognition biases in the search of sustainability. *Ambio*, 48(6), 605–618. <https://doi.org/10.1007/s13280-018-1100-5>
- Frank, R. H. (1987). If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience? *The American Economic Review*, 77(4), 593–604. <https://www.jstor.org/stable/1814533>
- Frederiks, E. R., Stenner, K., & Hobman, E. V. (2015). Household energy use: Applying behavioural economics to understand consumer decision-making and behaviour. *Renewable and Sustainable Energy Reviews*, 41, 1385–1394. <https://doi.org/10.1016/j.rser.2014.09.026>
- Garnett, C., Crane, D., Brown, J., Kaner, E., Beyer, F., Muirhead, C., Hickman, M., Redmore, J., de Vocht, F., Beard, E. & Michie, S. (2018). Reported theory use by digital interventions for hazardous and harmful alcohol consumption, and association with effectiveness: Meta-regression. *Journal of Medical Internet Research*, 20(2), e69. <https://doi.org/10.2196/jmir.8807>
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, 103(4), 650–669. <https://doi.org/10.1192/bjpo.bp.115.000224>
- Gigerenzer, G., & Selten, R. (2001). *Bounded Rationality The Adaptive Toolbox*. The MIT Press.
- Gotts, N. M., & Polhill, J. G. (2009). When and how to imitate your neighbours: Lessons from and for FEARLUS. *Journal of Artificial Societies and Social Simulation*, 12(3). <http://jasss.soc.surrey.ac.uk/12/3/2.html>
- Graybiel, A. M. (2008). Habits, rituals, and the evaluative brain. *Annual Review of Neuroscience*, 31, 359–387. <https://doi.org/10.1146/annurev.neuro.29.051605.112851>
- Gräbner, C., Bale, C. S. E., Furtado, B. A., Alvarez-Pereira, B., Gentile, J. E., Henderson, H., & Lipari, F. (2017). Getting the Best of Both Worlds? Developing Complementary Equation-Based and Agent-Based Models. *Computational Economics*, 53(2), 763–782. <https://doi.org/10.1007/s10614-017-9763-8>
- Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W. M., Railsback, S. F., Thulke, H.-H., Weiner, J., Wiegand, T., & DeAngelis, D. L. (2005). Pattern-oriented modeling of agent-based complex systems: lessons from ecology. *Science*, 310(5750), 987–991. <https://doi.org/10.1126/science.1116681>
- Grimm, V., Railsback, S. F., Vincenot, C. E., Berger, U., Gallagher, C., DeAngelis, D. L., Edmonds, B., Ge, J., Giske, J., Groeneveld, J., Johnston, A. S. A., Milles, A., Nabe-Nielsen, J., Polhill, J. G., Radchuk, V., Rohwäder, M. -S., Stillman, R. A., Thiele, J. C., & Ayllón, D. (2020). The ODD Protocol for describing Agent-Based and other simulation models: a second update to improve clarity, replication and structural realism. *Journal of Artificial Societies and Social Simulation* 23(2), 7 <http://jasss.soc.surrey.ac.uk/23/2/7.html>
- 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. *Environmental Modelling & Software*, 87, 39–48. <https://doi.org/10.1016/j.envsoft.2016.10.008>
- Hartig, F., Calabrese, J. M., Reineking, B., Wiegand, T., & Huth, A. (2011). Statistical inference for stochastic simulation models - theory and application. *Ecology Letters*, 14(8), 816–827. <https://doi.org/10.1111/j.1461-0248.2011.01640.x>
- Heckbert, S., Baynes, T., & Reeson, A. (2010). Agent-based modeling in ecological economics. *Annals of the New York Academy of Sciences*, 1185(1), 39–53. <https://doi.org/10.1111/j.1749-6632.2009.05286.x>
- Hewstone, M., Stroebe, W., & Jonas, K. (2015). *An Introduction to Social Psychology*. Wiley.
- Huber, R., Bakker, M., Balmann, A., Berger, T., Bithell, M., Brown, C., Grêt-Regamey, A., Xiong, H., Le, Q. B., Mack, G., Meyfroidt, P., Millington, J., Müller, B., Polhill, J. G., Sun, Z., Seid, R., Troost, C., & Finger, R. (2018). Representation of decision-making in European agricultural agent-based models. *Agricultural Systems*, 167, 143–160. <https://doi.org/10.1016/j.agry.2018.09.007>
- Jager, W. (2000). *Modelling consumer behaviour*. <https://www.rug.nl/research/portal/files/9914467/thesis.pdf>
- Jager, W. (2007). The four P's in social simulation, a perspective on how marketing could benefit from the use of social simulation. *Journal of Business Research*, 60(8), 868–875. <https://doi.org/10.1016/j.jbusres.2007.02.003>
- Jager, W. (2017). Enhancing the realism of simulation (EROS): On implementing and developing psychological theory in social simulation. *Journal of Artificial Societies & Social Simulation*, 20(3). <https://doi.org/10.18564/jasss.3522>

- Jager, W., Abramczuk, K., Komendant-Brodowska, A., Baczko-Dombi, A., Fecher, B., Sokolovska, N., Spits, T., (2020). Looking into the educational mirror: why computation is hardly being taught in the social sciences, and what to do about it. Proceedings of the Social Simulation Conference 2018, Mainz, Germany.
- Jager, W., & Ernst, A. (2017). Social Simulation in Environmental Psychology, Introduction of the special issue. *Journal of Environmental Psychology*, 52, 114–118. <https://doi.org/10.1016/j.jenvp.2017.07.002>
- 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. *Ecological Economics*, 35, 357–379. [https://doi.org/10.1016/S0921-8009\(00\)00220-2](https://doi.org/10.1016/S0921-8009(00)00220-2)
- Janssen, M. A. (2016). Impact of diverse behavioral theories on environmental management: explorations with Daisyworld. In T. M. K. Roeder, P. I. Frazier, R. Szechtman, E. Zhou, T. Huschka, & S. E. Chick (Eds.), *Proceedings of the 2016 Winter Simulation Conference*, IEEE Press. ISBN 978-1-5090-4484-9
- Janssen, M.A. (2017) The Practice of Archiving Model Code of Agent-Based Models, *Journal of Artificial Societies and Social Simulation* 20(1), 2. <http://jasss.soc.surrey.ac.uk/20/1/2.html> DOI: 10.18564/jasss.3317
- Janssen, M. A., & Baggio, J. A. (2017). Using agent-based models to compare behavioral theories on experimental data: Application for irrigation games. *Journal of Environmental Psychology*, 52, 194–203. <https://doi.org/10.1016/j.jenvp.2016.04.018>
- Kahneman, D. (2003). Maps of Bounded Rationality: Psychology for Behavioral Economics. *The American Economic Review*, 93(5), 1449–1475. <https://doi.org/10.1257/000282803322655392>
- Kahneman, D., & Tversky, A. (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica*, 47(2), 263–292. <https://doi.org/10.2307/1914185>
- Kiesling, E., Günther, M., Stummer, C., & Wakolbinger, L. M. (2012). Agent-based simulation of innovation diffusion: A review. *Central European Journal of Operations Research*, 20(2), 183–230. <https://doi.org/10.1007/s10100-011-0210-y>
- Klabunde, A., & Willekens, F. (2016). Decision-Making in Agent-Based Models of Migration: State of the Art and Challenges. *European Journal of Population*, 32(1), 73–97. <https://doi.org/10.1007/s10680-015-9362-0>
- Laatabi, A., Marilleau, N., Nguyen-Huu, T., Hbid, H., & Babram, M. A. (2018). ODD+2D: An ODD based protocol for mapping data to empirical ABMs. *Journal of Artificial Societies and Social Simulation*, 21(2), 9. <https://doi.org/10.18564/jasss.3646>
- Lempert, R. J., Groves, D. G., Popper, S. W., & Bankes, S. C. (2006). A general, analytic method for generating robust strategies and narrative scenarios. *Management Science*, 52(4), 514–528. <https://doi.org/10.1287/mnsc.1050.0472>
- Levine, J., Chan, K. M. A., & Satter, T. (2015). From rational actor to efficient complexity manager: Exorcising the ghost of Homo Economicus with a unified synthesis of cognition research. *Ecological Economics*, 114, 22–32. <https://doi.org/10.1016/j.ecolecon.2015.03.010>
- Lindenberg, S., & Steg, L. (2007). Normative, gain and hedonic goal frames guiding environmental behavior. *Journal of Social Issues*, 63(1), 117–137. <https://doi.org/10.1111/j.1540-4560.2007.00499.x>
- Maddux, J. E., & Rogers, R. W. (1983). Protection Motivation and Self-Efficacy: A Revised Theory of Fear Appeals and Attitude Change. *Journal of Experimental Social Psychology*, 19, 469–479. [https://doi.org/10.1016/0022-1031\(83\)90023-9](https://doi.org/10.1016/0022-1031(83)90023-9)
- Masterson, V. A., Stedman, R. C., Enqvist, J., Tengö, M., Giusti, M., Wahl, D., & Svedin, U. (2017). The contribution of sense of place to social-ecological systems research: A review and research agenda. *Ecology and Society*, 22(1), 49. <https://doi.org/10.5751/ES-08872-220149>
- Meyfroidt, P. (2013). Environmental cognitions, land change, and social–ecological feedbacks: an overview. *Journal of Land Use Science*, 8(3), 341–367. <https://doi.org/10.1080/1747423X.2012.667452>
- Milner-Gulland, E. J. (2012). Interactions between human behaviour and ecological systems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1586), 270–278. <https://doi.org/10.1098/rstb.2011.0175>
- Monroe, K. R. (2001). Paradigm Shift: From Rational Choice to Perspective. *International Political Science Review*, 22(2), 151–172. <https://doi.org/10.1177/0192512101222002>
- Moore, H. E., & Boldero, J. (2017). Designing Interventions that Last: A Classification of Environmental Behaviors in Relation to the Activities, Costs, and Effort Involved for Adoption and Maintenance. *Frontiers in Psychology*, 8, 1874. <https://www.frontiersin.org/article/10.3389/fpsyg.2017.01874>
- Muldoon, R., Liscandra, C., Bicchieri, C., Hartmann, S., & Sprenger, J. (2013). On the emergence of descriptive norms. *Politics, Philosophy & Economics*, 13(1), 3–22. <https://doi.org/10.1177/1470594X12447791>
- Muelder, H., & Filatova, T. (2018). One Theory - Many Formalizations: Testing Different Code Implementations of the Theory of Planned Behaviour in Energy Agent-Based Models. *Journal of Artificial Societies and Social Simulation*, 21(4), 5. <https://doi.org/10.18564/jasss.3855>
- Müller, B., Balbi, S., Buchmann, C. M., de Sousa, L., Dressler, G., Groeneveld, J., Klassert, C. J., Le, Q. B., Millington, J. D. A., Nolzen, H., Parker, D. C., Polhill, J. G., Schlüter, M., Schulze, J., Schwarz, N., Sun, Z., Taillandier, P., & Weise, H. (2014). Standardised and transparent model descriptions for agent-based models: Current status and prospects. *Environmental Modelling and Software*, 55, 156–163. <https://doi.org/10.1016/j.envsoft.2014.01.029>
- 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. *Environmental Modelling & Software*, 48, 37–48. <http://dx.doi.org/10.1016/j.envsoft.2013.06.003>
- Muhar, A., Raymond, C. M., van den Born, R. J. G., Bauer, N., Böck, K., Braito, M., Buijs, A., Flint, C., de Groot, W. T., Ives, C. D., Mitrofanenko, T., Pliening, T., Tucker, C., & van Riper, C. J. (2018). A model integrating social-cultural concepts

- of nature into frameworks of interaction between social and natural systems. *Journal of Environmental Planning and Management*, 61(5–6), 756–777. <https://doi.org/10.1080/09640568.2017.1327424>
- Nyborg, K., Anderies, J. M., Dannenberg, A., Lindahl, T., Schill, C., Schlüter, M., Adger, W. N., Arrow, K. J., Barrett, S., Carpenter, S., Chapin, F. S., Crépin, A.-S., Daily, G., Ehrlich, P., Folke, C., Jager, W., Kautsky, N., Levin, S. A., Madsen, O. J., Polasky, S., Scheffer, M., Walker, B., Weber, E. U., Wilen, J., Xepapadeas, A., & de Zeeuw, A. (2016). Social norms as solutions. Policies may influence large-scale behavioral tipping. *Science*, 354(6308), 42–43. <https://doi.org/10.1126/science.aaf8317>
- Nyblade, B., O'Mahony, A., & Sieck, K. (2019). Building on Social Science: Theoretic Foundations for Modelers. In P. K. Davis, A. O'Mahony, & J. Pfautz (Eds.), *Social-Behavioral Modeling for Complex Systems* (pp. 63–99). Wiley. <https://doi.org/10.1002/9781119485001.ch4>
- O'Hare, M., Plevin, R. J., Martin, J. I., Jones, A. D., Kendall, A., & Hopson, E. (2009). Proper accounting for time increases crop-based biofuels' greenhouse gas deficit versus petroleum. *Environmental Research Letters*, 4(2), 024001. <https://doi.org/10.1088/1748-9326/4/2/024001>
- Page, S. E. (1997). On Incentives and Updating in Agent Based Models. *Computational Economics*, 10, 67–87. <https://doi.org/10.1023/A:1008625524072>
- Pavlov, I. (1927). *Conditioned reflexes*. Oxford University Press, New York, NY, US.
- Peterson, M. (2017). *An Introduction to Decision Theory*, 2nd ed. Cambridge University Press.
- Poile, C., & Safayeni, F. (2016). Using computational modeling for building theory: A double edged sword. *Journal of Artificial Societies and Social Simulation*, 19(3). <https://doi.org/10.18564/jasss.3137>
- Pope, A. J., & Gimblett, R. (2015). Linking Bayesian and Agent-based Models to Simulate Complex Social-ecological Systems in Semi-arid Regions. *Frontiers in Environmental Science*, 3, 1–9. <https://doi.org/10.3389/fenvs.2015.00055>
- Rangoni, R., & Jager, W. (2017). Social dynamics of littering and adaptive cleaning strategies explored using agent-based modelling. *Journal of Artificial Societies and Social Simulation*, 20(2), 1. <https://doi.org/10.18564/jasss.3269>
- Risjord, M. (2019). Middle-range theories as models: New criteria for analysis and evaluation. *Nursing Philosophy*, 20, e12225. <https://doi.org/10.1111/nup.12225>
- Ritzer, G. (ed.) (2012). *Encyclopedia of Social Theory*. Sage.
- Rollins, N. D., Barton, C. M., Bergin, S., Janssen, M. A., & Lee, A. (2014). A computational model library for publishing model documentation and code. *Environmental Modelling and Software*, 61, 59–64. <https://doi.org/10.1016/j.envsoft.2014.06.022>
- Scarlett, L., Boyd, J., Brittain, A., Shabman, L., & Brennan, T. (2013). *Catalysts for Conservation: Exploring Behavioral Science Insights for Natural Resource Investments*. Resources for the future report. Retrieved from <https://www.rff.org/documents/443/RFF-Rpt-BehavioralScienceEconomicInsights.pdf>
- Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., Janssen, M. A., McAllister, R. R. J., Müller, B., Orach, K., Schwarz, N., & Wijermans, N. (2017). A framework for mapping and comparing behavioral theories in models of social-ecological systems. *Ecological Economics*, 131, 21–35. <https://doi.org/10.1016/j.ecolecon.2016.08.008>
- Schmolke, A., Thorbek, P., DeAngelis, D. L., & Grimm, V. (2010). Ecological models supporting environmental decision making: a strategy for the future. *Trends in Ecology & Evolution*, 25(8), 479–486. <https://doi.org/10.1016/j.tree.2010.05.001>
- Schulze, J., Müller, B., Groeneveld, J., & Grimm, V. (2017). Agent-Based Modelling of Social-Ecological Systems: Achievements, Challenges, and a Way Forward. *Journal of Artificial Societies and Social Simulation*, 20(2), 8. <https://doi.org/10.18564/jasss.3423>
- Silverman, E., Bijak, J., & Noble, J. (2011). Feeding the beast: Can computational demographic models free us from the tyranny of data? *European Conference on Artificial Life*, 276, 747–754. Retrieved from <http://eprints.ecs.soton.ac.uk/22839/>
- Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118. <https://doi.org/10.2307/1884852>
- Simon, H. A. (1978). Rationality as Process and as Product of Thought. *The American Economic Review*, 68(2), 1–16. <https://doi.org/10.1017/CBO9780511598951.005>
- Skinner, B.F. (1938). *The behavior of organisms*. Appleton-Century.
- Smajgl, A., & Barreteau, O. (2017). Framing options for characterising and parameterising human agents in empirical ABM. *Environmental Modelling and Software*, 93, 29–41. <https://doi.org/10.1016/j.envsoft.2017.02.011>
- Smith, E. R., & Conrey, F. R. (2007). Agent-based modeling: A new approach for theory building in social psychology. *Personality and Social Psychology Review*, 11(1), 87–104. <https://doi.org/10.1177/1088868306294789>
- Speelman, E.N., (2014). *Gaming and simulation to explore resilience of contested agricultural landscapes*. Doctoral Thesis, Wageningen University, Farming Systems Ecology
- Steg, L., van den Berg, A. E., & de Groot, J. I. M. (2012). *Environmental Psychology: An Introduction*. Wiley-Blackwell.
- Steg, L., & Vlek, C. (2009). Encouraging pro-environmental behaviour: An integrative review and research agenda. *Journal of Environmental Psychology*, 29(3), 309–317. <https://doi.org/10.1016/j.jenvp.2008.10.004>
- Sullivan, A. (2002). Bourdieu and education: How useful is Bourdieu's theory for researchers? *The Netherlands' Journal of Social Sciences*, 38(2), 144–166.
- Thiele, J. C., Kurth, W., & Grimm, V. (2014). Facilitating Parameter Estimation and Sensitivity Analysis of Agent-Based Models: A Cookbook Using NetLogo and R. *Journal of Artificial Societies and Social Simulation*, 17(3), 11. <https://doi.org/10.18564/jasss.2503>



- Turner, B.S., Kyung-Sup, C., Epstein, C.F., Kivisto, P., Outhwaite, W., Ryan, J.M. (eds., 2017). *The Wiley Blackwell Encyclopedia of Social Theory*. Wiley-Blackwell.
- van Vugt, M., Griskevicius, V., & Schultz, P. W. (2014). Naturally Green: Harnessing Stone Age Psychological Biases to Foster Environmental Behavior. *Social Issues and Policy Review*, 8(1), 1–32. <https://doi.org/10.1111/sipr.12000>
- Watson, A. J., & Lovelock, J. E. (1983). Biological homeostasis of the global environment: the parable of Daisyworld. *Tellus B*, 35 B(4), 284–289. <https://doi.org/10.1111/j.1600-0889.1983.tb00031.x>
- West, R., Godinho, C. A., Bohlen, L. C., Carey, R. N., Hastings, J., Lefevre, C. E., & Michie, S. (2019). Development of a formal system for representing behaviour-change theories. *Nature Human Behaviour*, 3, 526-536. <https://doi.org/10.1038/s41562-019-0561-2>
- Wijermans, N., Boonstra, W. J., Orach, K., Hentati-Sundberg, J., & Schlüter, M. (2020). Behavioural diversity in fishing—Towards a next generation of fishery models. *Fish and Fisheries*, 274(2), 30–19. <http://doi.org/10.1111/faf.12466>
- Wilensky, U. (1999). NetLogo. <http://ccl.northwestern.edu/netlogo/>. Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL.
- World Bank (2015). *World Development Report 2015: Mind, Society, and Behavior*. Retrieved from <http://www.worldbank.org/en/publication/wdr2015>
- Yu, Q., Verburg, P. H., & Wu, W. (2019). Environmental cognitions mediate the causal explanation of land change. *Journal of Land Use Science*, 13(5), 535–548. <https://doi.org/10.1080/1747423X.2019.1567837>