Springer Proceedings in Complexity

Harko Verhagen
Melania Borit
Giangiacomo Bravo
Nanda Wijermans *Editors*

Advances in Social Simulation

Looking in the Mirror

Springer

Springer Proceedings in Complexity

Springer Proceedings in Complexity publishes proceedings from scholarly meetings on all topics relating to the interdisciplinary studies of complex systems science. Springer welcomes book ideas from authors. The series is indexed in Scopus.

Proposals must include the following:

- name, place and date of the scientific meeting
- a link to the committees (local organization, international advisors etc.)
- scientific description of the meeting
- list of invited/plenary speakers
- an estimate of the planned proceedings book parameters (number of pages/articles, requested number of bulk copies, submission deadline)

Submit your proposals to: christoph.baumann@springer.com

More information about this series at http://www.springer.com/series/11637

Harko Verhagen • Melania Borit
Giangiacomo Bravo • Nanda Wijermans
Editors

Advances in Social Simulation

Looking in the Mirror

Springer

Editors
Harko Verhagen
Department of Computer and Systems Sciences
Stockholm University
Kista, Sweden

Giangiacomo Bravo
Linnaeus University Centre for Data Intensive Sciences and Applications
Växjö, Sweden

Melania Borit
Norwegian College of Fishery Science
UiT The Arctic University of Norway
Tromsø, Norway

Nanda Wijermans
Stockholm Resilience Centre
Stockholm University
Stockholm, Sweden

ISSN 2213-8684 ISSN 2213-8692 (electronic)
Springer Proceedings in Complexity
ISBN 978-3-030-34126-8 ISBN 978-3-030-34127-5 (eBook)
https://doi.org/10.1007/978-3-030-34127-5

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Since the first larger scale conferences in the early 1990s, social simulation has rapidly evolved into becoming a methodology known by researchers and practitioners in different areas including sociology, political science, economics, history, and ecology, to name a few. Each of these disciplines has to address what sociologists call the micro-macro link – how individual decision-making entities interact to produce effects at an aggregate level (*emergence*) while, at the same time, aggregate-level mechanisms and processes affect, enable, and limit the choices and/or behaviors of the individuals (*immergence*). Addressing these dynamic processes using social simulation, in general, and agent-based modeling, in particular, has been proven successful for both research and application in, for instance, policy-making. The ability to zoom in on interacting individuals and zoom out to catch aggregate-level processes and patterns, while still able to systematically experiment with different parameters, is the key advantage of simulation over more analytical approaches. The social simulation conferences have been instrumental in the development of methodological and epistemological contributions by enabling a constructive and meaningful community, allowing for discussions with colleagues from different research areas and backgrounds. These conferences are at the heart of the European Social Simulation Association (ESSA), founded in 2003 as a scientific society to encourage the advancement of social simulation research, to promote cooperation among scientists and scholars and practitioners from different disciplines using computational models and methodologies to advance the understanding of social processes, and to encourage and support the development in Europe of education in social simulation.

This book reflects the recent developments as presented and discussed at the 14th Social Simulation Conference in Stockholm, Sweden (20–24 August 2018), hosted by the Department of Computer and Systems Sciences of Stockholm University and co-organized together with the UiT –The Arctic University of Norway, Tromsø, Norway, Stockholm Resilience Centre of Stockholm University, and Linnaeus University, Växjö, Sweden. The theme of the conference was "Looking in the Mirror," reflecting adolescence and the coming of age of the conference and of ESSA itself. The conference actively stimulated discussions every day by starting with keynotes and closing with daily reflection conversations with selected participants. The stimulating reflections of our keynotes are presented in the first two chapters of the proceedings, reflecting from an inside and outside perspective on the development of the social simulation community. Bruce Edmonds, the community critical voice from its very beginnings, reflected on the importance of context to understand rarely addressed social phenomena (see Chapter 1). Milena Tsvetkova, a computational social scientist, shared her work on large-scale online experiments, showing the breadth of computational social science research beyond social simulation. Finally, Julie Zahle, philosopher of social sciences, argued about agent-based computational modeling and its relationship to different strands of methodological individualism, forcing social simulation modelers to reflect on whether they actually do what they claim to do (see Chapter 2).

The regular contributions were organized in a number of different tracks corresponding to different foci and disciplines within the community. In total, the conference received 110 contributions consisting of 35 long papers, 59 short papers, and 16 posters. All contributions received at least 2 reviews in a single-blind process, and 109 contributions were accepted, with some contributions referred to another category. Apart from the 49 papers (including 2 invited speaker papers) presented in this volume, a set of papers will appear in a special issue of the *Journal of Artificial Societies and Social Simulation* (*JASSS*).

The areas represented were as follows:

Best practices and new methodologies of representing dynamic networks with a focus on archaeological applications (Chaps. 27 and 40)

Models that explicitly address the social, relational side of life, efforts to make the choices and assumptions underlying agent behavior more explicit (Chaps. 13, 16, and 42)

The use of social simulation and computational social science in solving contemporary challenges; new, forgotten, or underrepresented methodologies and application areas; and the future of the research area and its community within the broader social sciences and computing domains (Chaps. 38 and 47)

Models that elaborate on the representation of agent decision-making (Chaps. 34 and 49)

Modeling social science aspects of socio-ecological systems (Chapter 48)

Presentation and discussion of recent research findings which challenges economic models by integrating empirically sound assumptions and integrating findings from other disciplines to develop economic theory (Chaps. 6, 11, 18, 21, and 46)

Simulation models of organizations as complex adaptive systems where the processes of organizing are linked to factors such as structural interdependencies among organizational subunits and individual behavior (Chapter 26)

Best practices in courses and/or knowledge transfer related to social simulation (Chaps. 8, 22, 37, 43, and 45)

How to address important challenges that need to be addressed, such as scaling issues, counterfactual scenery assessments, external/internal validities, coding efficiencies, isolation mechanisms, understandability of complex models, etc., empirically grounded simulations that aim to design and evaluate policy interventions by building more realistic models that minimize possible unintentional consequences due to unrealistic assumptions, and the need to share our work and collaborate on large-scale complex models (Chaps. 10, 11 and 33)

Simulation models that study these challenges with a social-ecological complex systems lens and a focus on the social behavior components of the system (Chaps. 15, 19, 30, 31, and 32)

The interplay between social simulation and games (Chaps. 23, 24)

Opportunities and hindrances for use of social simulation with policy actors (Chaps. 3, 7, 9, 17, 28, 35, and 44)

Integration of qualitative evidence in simulation models in a systematic way (Chaps. 4, 29, 36, 38, and 41)

Validation challenges for agent-based modelers (Chaps. 5, 12, and 14)

We wish to acknowledge that while Stockholm University supported the organization by providing the conference venue and website, the UiT – The Arctic University of Norway and ESSA sponsored scholarships for early-stage researchers.

Organization

General Chair

Harko Verhagen	Stockholm University, Sweden

Program Committee Chairs

Melania Borit	UiT The Arctic University of Norway, Norway
Giangiacomo Bravo	Linnaeus University, Sweden
Nanda Wijermans	Stockholm University, Sweden

Program Committee

Diana Francisca Adamatti	Universidade Federal do Rio Grande, Brazil
Petra Ahrweiler	Johannes Gutenberg University Mainz, Germany
Fred Amblard	Université Toulouse 1 Capitole, France
Pedro Ribeiro De Andrade	INPE, Brazil
Luis Antunes	University of Lisbon, Portugal

Jennifer Badham	Queen's University Belfast, Ireland
Franco Bagnoli	University of Florence, Italy
Albert Bakhtizin	Central Economics and Mathematics Institute of Russian Academy of Sciences, Russia
Stefano Balietti	Microsoft, USA
Tina Balke	University of Surrey, UK
Peter Barbrook-Johnson	University of Surrey, UK
Galina Besstremyannaya	Stanford University, USA
Federico Bianchi	University of Brescia, Italy
Mike Bithell	University of Cambridge, UK
Leonardo Bocchi	University of Firenze, Italy
Riccardo Boero	Los Alamos National Laboratory, USA
Melania Borit	UiT The Arctic University of Norway, Norway
Giangiacomo Bravo	Linnaeus University, Sweden
Emile Chappin	Delft University of Technology, the Netherlands
Edmund Chattoe-Brown	University of Leicester, UK
Shu-Heng Chen	National Chengchi University, Taiwan
S Anders Christensson	Swedish National Defence University, Sweden
Nick Collier	Argonne National Laboratory, USA
Andrew Crooks	George Mason University, USA
Koen de Koning	University of Twente, the Netherlands
Guillaume Deffuant	Cemagref, France
Gennaro Di Tosto	The Ohio State University, USA
Frank Dignum	Utrecht University, the Netherlands
Virginia Dignum	Delft University of Technology, the Netherlands
Ozge Dilaver	University of Surrey, UK
Stefano Duca	ETH Zurich, Switzerland
Julia Eberlen	Belgian National Fund for Scientific Research, Belgium
Bruce Edmonds	Manchester Metropolitan University, UK
Corinna Elsenbroich	University of Surrey, UK
Giorgio Fagiolo	Sant'Anna School of Advanced Studies
Mike Farjam	Linnaeus University, Sweden
Thomas Feliciani	University of Groningen, the Netherlands
Andreas Flache	University of Groningen, the Netherlands
Cesar Garcia-Diaz	Universidad de los Andes, Colombia
Armando Geller	Scensei, Switzerland
Amineh Ghorbani	Delft University of Technology, the Netherlands
Francesca Giardini	Groningen University, the Netherlands
Nigel Gilbert	University of Surrey, UK
Gustavo Giménez-Lugo	Federal University of Technology, Paraná, Brazil

Yusuke Goto	Iwate Prefectural University, Japan
Nick Gotts	UK
William Griffin	Arizona State University, USA
Francisco Grimaldo	Universitat de València, Spain
Andrea Guazzini	University of Florence, Italy
Laszlo Gulyas	Eötvös Loránd University, Hungary
Mirsad Hadzikadic	UNC Charlotte, USA
David Hales	The Open University, UK
Rainer Hegselmann	Bayreuth University, Germany
Reiko Hishiyama	Waseda University, Japan
Gert Jan Hofstede	Wageningen University, the Netherlands
Segismundo Izquierdo	Universidad de Valladolid, Spain
Wander Jager	University of Groningen, the Netherlands
Marco Janssen	Arizona State University, USA
Bogumil Kaminski	Warsaw School of Economics, Poland
Toshiyuki Kaneda	Nagoya Institute of Technology, Japan
Hajime Kita	Kyoto University, Japan
Andreas Koch	University of Salzburg, Austria
Matthew Koehler	The MITRE Corporation, USA
Setsuya Kurahashi	University of Tsukuba, Japan
Stephan Leitner	Alpen-Adria Universitaet Klagenfurt, Germany
Zoran Levnajic	Faculty of Information Studies in Novo mesto, Slovenia
Emilie Lindkvist	Stockholm University, Sweden
Emiliano Lorini	IRIT, France
Iris Lorscheid	University of Applied Sciences Europe, Germany
André Martins	University of São Paulo, Brazil
Ruth Meyer	Manchester Metropolitan University, UK
Francisco José Miguel Quesada	Universitat Autònoma de Barcelona, Spain
Makoto Mizuno	Meiji University, Japan
John Murphy	Argonne National Laboratory, USA
Selcan Mutgan	Linköping University
Michael Mäs	ETH Zurich, Switzerland
Luis Gustavo Nardin	University of Idaho, USA
Pablo Noriega	Artificial Intelligence Research Institute (IIIA), Spain
Emma Norling	The University of Sheffield, UK
Kotaro Ohori	Fujitsu Laboratories Ltd., Japan
Isamu Okada	Soka University, Japan
Jonathan Ozik	Argonne National Laboratory and The University of Chicago, USA

Mario Paolucci	Institute of Cognitive Sciences and Technologies, Italy
Nicolas Payette	Université du Québec à Montréal, Canada
Diane Payne	University College Dublin, Ireland
Denis Phan	GEMAS CNRS & Université Paris IV Sorbonne, France
Jakub Piskorski	Polish Academy of Sciences, Poland
Gary Polhill	The James Hutton Institute, UK
Kavin Preethi Narasimhan	University of Surrey, UK
Andreas Pyka	Universität Hohenheim, Germany
William Rand	University of Maryland, USA
Juliette Rouchier	CNRS-LAMSADE, France
Jordi Sabater Mir	IIIA-CSIC, Spain
Anxo Sanchez	Universidad Carlos III de Madrid, Spain
Jorge Santos	UiT NFH, Norway
Aki-Hiro Sato	Kyoto University, Japan
Frederik Schaff	FernUni Hagen, Germany
Maja Schlüter	Stockholm University, Sweden
Geeske Scholz	University of Osnabrück, Germany
Roman Seidl	Institute for Applied Ecology, Germany
Jaime Sichman	University of São Paulo, Brazil
Alex Smajgl	Mekong Region Futures Institute, Thailand
Erika Speelman	Wageningen University, the Netherlands
Flaminio Squazzoni	University of Brescia, Italy
Przemysław Szufel	Warsaw School of Economics
Chung-Ching Tai	Department of Economics, Tunghai University, Taiwan
Karoly Takacs	Hungarian Academy of Sciences, Centre for Social Sciences
Takao Terano	The University of Tokyo
Pietro Terna	University of Torino, Italy
Klaus G. Troitzsch	Germany
Loïs Vanhée	Université de Caen Normandie, France
Harko Verhagen	Stockholm University, Sweden
Daniele Vilone	Istituto di Scienza e Tecnologia della Cognizione (ISTC) – CNR, Italy
Annie Waldherr	University of Muenster, Germany
Friederike Wall	Alpen-Adria-Universitaet Klagenfurt, Germany
Wybo Wiersma	University of Oxford, UK
Nanda Wijermans	Stockholm University, Sweden
Ryuichi Yamamoto	Waseda University, Japan

Additional Reviewers

Molood Ale Ebrahim
Koen Degeling
Arthur Feinberg
Christian Klöckner
Friedrich Krebs
Quang Bao Le
Nina Schwarz

Kista, Sweden — Harko Verhagen
Tromsø, Norway — Melania Borit
Växjö, Sweden — Giangiacomo Bravo
Stockholm, Sweden — Nanda Wijermans
July 2019

Contents

Chapter 1
How Social Simulation Could Help Social Science Deal with Context

Bruce Edmonds

Abstract Much human cognition and behaviour is context-sensitive, but context (especially social context) has largely not been explicitly represented or included in theories of explanations of behaviour. Some of this is due to the fact that the word "context" is over-used and so has a variety of subtly different meanings but more to do with the perceived difficulties of dealing with context. Quantitative social science has tended to ignore context, treating contextual variation as "noise". Qualitative social science has often almost deified context, resisting any attempts to generalise from specific contexts. This paper suggests that agent-based modelling could play a key role in dealing with context, representing it, understanding it and thus allowing the well-founded integration of qualitative and quantitative evidence.

Keywords Context · Social science · Social simulation · Quantitative · Qualitative · Agent-based modelling

1.1 Introduction

That context is important for understanding social phenomena should be uncontroversial; yet dealing with it has been largely avoided. On one side, quantitative social scientists tend to fit data that originates from a variety of contexts with a single model (e.g. variants of linear regression) on the grounds that they are only interested in generic patterns. At the other extreme, qualitative researchers interested in rich ("thick") descriptions of observations and experience take context seriously, ensuring that they take care to describe relevant aspects of the context in what they record and discuss but tend to resist any generalisations that cross contexts. The point is that a crucial issue is not being explicitly addressed: that of context-dependency itself.

B. Edmonds (✉)
Centre for Policy Modelling, Manchester Metropolitan University, Manchester, UK
e-mail: bruce@edmonds.name

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_1

Let us start by making clear what everybody knows: people behave differently in different *kinds* of situation that we can effectively recognise these kinds of situation and use them to understand, and even predict, what people will do in these situations. For example, we all recognise a lecture and know the social norms, habits, conventions, roles, etc. that pertain there. If the lecture is declared finished and coffee or wine served to celebrate something, the context has changed and everybody will behave differently. To take another example, traders in a stock market behave very differently during a bull and bear market (e.g. [11]). In a bull market, it is relatively easy to make some money and traders might seek to maximise their profits and endure quite high risk. In a bear market, traders are in danger of losing their job, so it might be more important to not be the worst performer in their group above all. In both cases, understanding behaviour is much easier and more effective if you divide the case into the different contexts. To produce models of behaviour that pertain to both lecture and celebration or to both bear and bull markets involves a much more complex, ineffective and abstract model. So why don't quantitative social scientists pay any attention to this common sense knowledge? Similarly, we all are able to recognise the difference between a lecture and a celebration or traders between a bull and a bear market and, without thinking about it much, apply the appropriate knowledge to each. So why don't some qualitative social scientists accept the usefulness of cautious generalisation over particular situations?

This paper seeks to examine the difficulty of talking about and studying social phenomena in a way that includes context but also suggests some ways forward to do this. *Firstly*, there is some discussion of the different ways that the word "context" is used, distinguishing between some different uses and trying to make clear what I intend by the word. *Then*, the paper looks at and critiques some of the ways that social science deals with (or avoids) the issue of context, looking at the responses from quantitative and then qualitative approaches. *Thirdly*, it looks at two ways in which agent-based social simulation could represent or be informed by social context – via the use of context-sensitive cognitive models for agents and by using the analysis of narrative data (including information about the relevant social contexts) to help inform the specification of simulations. It ends with a plea to change how we do social science.

1.2 Talking About "Context"

Before we can look at the issues, there are many potential confusions that can arise from the term "context". Thus I start by discussing this confusion and distinguishing between some different uses – making connections between the different meanings and trying to make my intent clear.

"Context" is a tricky word to use and a tricky phenomenon to pin down. Like other notorious c-words ("complexity" and "creativity"), it is often used as a "dustbin" concept – what one evokes when one's normal explanation fails. It can also be used as a "flag" to indicate that the research is qualitative since context is emphasised (almost deified) in the qualitative social sciences yet downplayed (usually ignored completely) in the quantitative social sciences. Context, as a word, is often used informally and hence has lots of subtly different usages, e.g. as documented in Hayes [9]. Finally it is not clear that a particular context can be reliably reified and talked about as *thing* at all. These difficulties may explain the reluctance of researchers to engage with context, knowing it is a notoriously slippery and difficult subject – I guess that many think it is better to avoid the swamp and only play on firmer ground. However, with a little care, I argue that the idea and its manifestations can be sensibly and usefully dealt with and the potential pay-off for the social sciences could be considerable.

1.2.1 Situational Context

Firstly "the context" can refer to the situation one is in [3] – for them the exterior situation *is* the context. This could be indicated by the exact coordinates and time; however this is not a very helpful notion. The details that could be potentially relevant to any series of events in any situation are indefinitely extensive. Rather it is usual to abstract from specific situations to *kinds* of situation, for example, going home on the train or shopping in a supermarket. The question "What was the context?" implies that the speaker does not have enough information about the situation some utterance or text comes from to understand it. The answer to such a question would not be to specify the precise situation but to give enough information about it to characterise the *kind* of situation one was in (e.g. "I was talking on the phone to my mother").

1.2.2 Cognitive Context

The fact that we can give enough information in a few words for the recipient to be able to infer the right kind of situation indicates that such recognition is not only feasible but also normal. It is well established that many aspects of human cognition are highly context-dependent, including memory, preferences, language use, language comprehension, decision-making, perception and reasoning [12, 16]. This implies that the brain has learned to reliably recognise these kinds of situation and effectively the same kinds as others do. The cognitive correlate of the kind of situation is called the "cognitive context" [9][1]. Though most of us, as individuals, do this unconsciously and with great facility (at least after childhood), we do not

[1]However "internal" factors such as emotion and current goals may also be inputs to determining this.

know how the brain does this, and it may be that it is very hard to replicate this recognition explicitly[2]. However this ability allows for the following heuristics: to learn knowledge with respect to the cognitive context currently being recognised and give preferential access to that knowledge when the same cognitive context occurs again. Thus when we enter a lecture, we do not have to "sift" through all the social norms we have learned for the relevant one, but those relevant automatically come to mind in that situation.

1.2.3 Social Context

Although cognitive context may be infeasible to determine in many cases, there is one case where this may be much easier – that where the context has been co-determined by many individuals in a society so that everybody recognises the same kinds of situation. Examples include the lecture, a celebration, commuting within a shared vehicle, religious ceremonies and an interview. Over time, specific norms, habits, language, spaces, technologies and even clothing might have been developed for that kind of situation, allowing the particular context to be easily distinguished. Of course, the reverse also happens: the more easily a particular context is distinguished, the more easily we will recognise it and develop specific practices, technologies and methods of coordination for it. Thus, over time, some contexts can become socially entrenched, acquire their own labels and be explicitly talked about. For this reason, such "social contexts" are much easier to identify and study than context in general.

Such social contexts can be very important since they allow for very different systems for social coordination to be developed for different kinds of situation. For example, how one coordinates behaviour on a fishing boat during periods of calm might be very different from that when a storm is approaching. It is not just the parameters of the coordination that change but the whole process.

Due to the difficulties involved in studying context, and a simple wish to avoid the extra complexity that they imply, researchers have tended to, in effect, avoid dealing with context head-on. A number of common research strategies have this effect. These will now be discussed in turn.

1.3 How Social Science Deals with Context

Here I briefly review some of the ways in which social science currently deals with (or avoids) context to motivate the need for the suggestions that follow in the next section. This is by no means a comprehensive survey, since that would far too

[2]Thus it may not make sense to assume that "the context" can always be reified as a distinct object that can be referred to, though as we argue it sometimes can be.

lengthy, but I hope this is sufficient to make the huge lacuna real for the reader. It includes two major approaches/assumptions made in quantitative social science and a critique of how some qualitative social science avoids the problem of context-dependency by leaving any cross-context generalisation implicit.

1.3.1 Quantitative Social Science

1.3.1.1 Context-Dependency and Randomness

For those researchers who claim to be “only interested in what behaviour is generic”, they may choose a model and “fit” it to some data, to see how good a fit it is (or whether it has a better fit than an alternative model). The variation not captured by the models is then attributed to “noise”, which is usually represented as some kind of randomness (sometimes this is indicated by the presence of an “error” term in equations). Typically the same, relatively simple model is fitted to the whole available data set and the extent that it “explains” the data – the likelihood that this fit is not by chance assessed (the so-called significance tests). In the social sciences, these are often variants of linear correlation models, though others variants also exist, such as the use of the “POMDP” class of models [10] in natural language processing.

The problem with this approach is that the generalisation could be occurring over different kinds of situation where different kinds of strategies might be being exhibited. The generic model “averages” over these different kinds of behaviour – producing a composite behaviour that might have elements of all of them, but misses some of the essential structural information about the observed.

A simple abstract example can illustrate this problem. Say there are two kinds of situation that occur within a sample of data: type **A** and type **B**. Within type **A**, variable *a* is strongly correlated with outcome *x*, and variable *b* is weakly anti-correlated with outcome *y*. Within type **B,** the opposite occurs: variable *a* is weakly anti-correlated with outcome *x*, and variable *b* is strongly correlated with outcome *y*. If types **A** and **B** occurred with roughly equal frequency, a generic correlation model relating *a* and *b* to *x* and *y* fitted to the complete data set might come to the conclusion that *a* is weakly correlated with *x* and *b* is weakly correlated with *y* for the whole domain at a significant level. This is illustrated below in Fig. 1.1. In this way, a lot of valuable information has been lost compared to a composite model that fitted separate models for each type. Comparing the generic model to the composite model, one would find that the generic model is not as strong and it misses the fact that there are parts of the data (from a specific kind of situation) with anti-correlations. Even an approach which included a variable to say whether a point belonged to type A or B would not help unless it was able to “switch” off and on the appropriate parts of the generic model.

If one imagines fitting a generic model over a great *many* kinds of situation, the expected result would be that many variables would be correlated with many

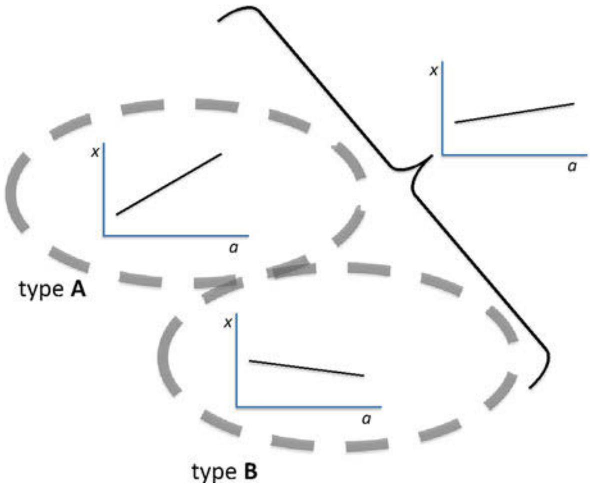


Fig. 1.1 An illustration of averaging out context-specific trends into a generic model

others at a significant level but only explaining a relatively small level of the total variation. This is, indeed, the result of many exercises in social science that apply generic models to data that may cover many different kinds of situation.

Not only has a lot of information been lost, but any policy based on such an analysis might be ineffective or even counter-productive for sub-groups of the population. Consider the case where there were twice as many of type **A** than of type **B**; then a generic correlation model fitted to the data might be that variable *a* is weakly correlated with *x* but there is no overall correlation of *b* with *y*. If the objective of policy is to increase *x* and *y*, then the inferred strategy would be to increase *a* only – despite the fact that this would have the contrary impact in a third of the cases. It seems obvious that if there were a technique to detect that there were essentially two different groups and to determine what model of behaviour fitted each, this might allow for a finer-grained understanding of data that might allow the more effective targeting of policy. Of course if it turned out that there was a substantial commonality between the separate models inferred, it might make sense to combine them into a generic model.

1.3.1.2 Over-generic Cognitive Models

Many of those that produce models based upon some kind of micro-specification simply *assume* that there is some generic model of behaviour that is valid across different contexts. The idea seems to be that there *must* be some generic cognitive model, albeit complex, that changes when the input to that agent or unit changes. I call the assumption that there must be a generic underlying model "behavioural foundationalism".

For example, neo-classical decision theory reduces all decisions between choices to a single model: that of a utility comparison on the consequences of decisions. More complex or social decisions are implemented by a more complex utility function. However, such an approach excludes any examination of the *process* by which decisions might be made[3], which might well be different in different circumstances – processes that might have very different collective outcomes from each other. For example, even if a process of individual consideration of the options and a social one (maybe looking what others are doing and imitating those who are most successful) might have a similar individual outcome for that individual, they might have very different collective outcomes. If many are following the imitative strategy, then there will be spreading waves of innovation. Another example of where the only behaviour modelled is that of social norms (e.g. [7])[4]. On a more mundane level, most of the cognitive algorithms within social simulations tend to be fairly simple and include no context-dependency at all, so in a way these all build in the assumption that a single simple algorithm will be sufficient to cover the different circumstances the agents face.

Whilst it is true that in some ultimate, biological sense, humans do have roughly the same equipment for making decisions – the nervous system – it is also true that this equipment takes years of external input, training, before it is very useful for making decisions. This suggests that a generic model of decision-making would have to be similarly complex and able to learn different strategies for different kinds of situation. There is ample evidence that many aspects of human cognition are context-dependent, including memory, decision-making, language, preferences and perception [12, 16]. One suspects that neo-classical economists simply hoped that they could produce physics-like models and thus bypass the complex and messy ways people actually make decisions, enabling them to find a shortcut that dealt with analytically modellable processes. However it has not had good empirical success. There are now so many exceptions to the received pattern of economic rationality that we should start to question whether this is the right starting point or whether this dogma might be better ditched.

It needs to be pointed out that many doing agent-based simulations are just as guilty of assuming a simple generic model of behaviour as neo-classical economists. Again this seems to be justified by an assumption that a more abstract model will be more general[5]. This is no more than a convenient hope. Whilst it might well be true that adding empirically based detail into a model might make it less general, the reverse does not work – simplifying away detail does not mean one achieves greater generality. The reason for this is clear – when simplifying, one does not know a priori what detail can be safely abstracted away. If one removes something essential to the phenomena being studied, then the result is a model that does not

[3] Simon's [15] distinction between procedural and substantive rationality

[4] Even social norms work in a complex and context-sensitive way [20].

[5] Although one suspects that often the reasons are more based on the mundane constraints of time and complication.

work for *any* observed cases, i.e. with no generality at all. Imagine abstracting away the variables from a linear model and just leaving the constant; this has not resulted in a more general model but one that is true almost nowhere.

A problem here is that an abstract model may often *seem* to be potentially applicable to a wide range of cases, but not in a precise manner. Here the model is used as a kind of analogy – that is, what the model parts refer to is not precisely defined but is left to each interpreter to construct "on the fly" – each person will interpret it in a different way. This is in contrast to a model where its relationship to what we might observe (represented by data of some kind) is well defined. Analogies provide very useful ways of thinking about a situation but do not give reliable or testable knowledge. Their success as an analogy does not give any guarantees that a more concrete version will be able to establish a more direct relationship with anything observable – developing an analogical model does not necessarily lead to an empirically validated one. In particular, a more abstract model of behaviour that appears to have general applicability (because it is used as an analogy) may well turn out to have less scope than one that is specific to a particular kind of situation in an empirically precise manner.

It may turn out that some elements of our behaviour can be understood in a generic manner, independent of the context, but this is something that needs to be demonstrated rather than assumed because it makes our job (as researchers) easier.

1.3.2 Qualitative Social Science

In contrast to the above approaches, many qualitative approaches pay a lot of attention to context – this is often described and included in their accounts and is by no means an afterthought or avoided. Indeed context is often deemed so important in these studies, that *any* possibility of generalisation to a different context is avoided – each context is unique. Thus one might have some high-quality observational or ethnographic work describing individual behaviour and strategies within a specific context but without any indication as to what could be learnt from this that might be useful elsewhere[6] – generalisation is often left to the reader here.

This is a highly defensible stance since generalisations are risky and open to criticism by others. By keeping to discussion of phenomena only *within* specific contexts, one can counter any objection with regard to the unique circumstances within the observed contexts – contexts that the presenting researcher has unique

[6]The exception is negative knowledge – counter examples to established assumptions – but this leads to the conclusion that we know nothing except specifics.

access to[7]. Here we have to opposite problem to overgeneralisation, to a situation where almost nothing is generalised at all[8].

In order for any knowledge to be useful, one needs to have some idea as to when it is applicable. Thus although detailed qualitative observations can expand our ideas of what people do in different situations – the possibilities – to be useful, we also need to know something about to what kinds of situation we can apply this knowledge.

1.4 Some Ways Agent-Based Social Simulation Could Deal with Context

Despite the difficulty of the subject and the corresponding circumvention in much social science, there are a number of approaches whereby we might at least start to touch upon context within social phenomena. These include using machine learning algorithms to attempt to infer context from suitably rich and extensive data [4, 17] and context-sensitive visualisation approaches, staging modelling to make the scope of sub-models clearer and extending qualitative elicitation techniques to better clarify implicit indications of context. However, here I only concentrate on two approaches which involve social simulation.

1.4.1 Implementing Context-Sensitive Agents in Social Simulations

Whilst it is very hard to include context-dependency within analytically solvable models, there is no reason why this needs to be the case with agent-based simulation models. However this does require a bit more "cognitive" machinery. Instead of each agent having a fixed resource of knowledge or behavioural rules, it needs to have different pools of such resources that can be selected depending on the context. In other words, the memory of the agent needs to be context-sensitive, so that context-relevant knowledge and behaviours can be preferentially applied in decision-making. Although this requires some technical changes, this is quite possible to do, ending up with an architecture as illustrated in Fig. 1.2. This sort

[7]Even if others have observed the same general kind of situation, it can always be claimed that this was at a different time or involved different actors with different goals – the defensive strategy that says every context is unique.

[8]To be precise, specific observations might be accompanied by imprecise and analogical discussion, but this is also immune to being wrong (except in maybe missing out a favorite dimension of a reader) due to its informality. Also it does not help in the identification of context or any other indications of when knowledge can be reliably used elsewhere.

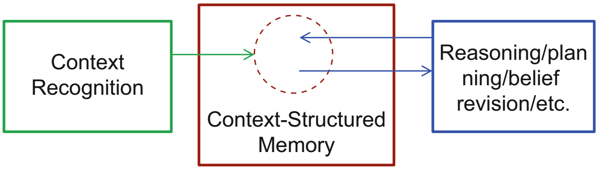


Fig. 1.2 The basic context-sensitive architecture for agents

of architecture also has major advantages in terms of the feasibility of learning or reasoning, since each of these is restricted to the relevant set of knowledge for that context. It also allows a well-structured integration between context recognition (which may leverage "fuzzy" machine learning techniques) with reasoning and belief update algorithms (which tend to be crisp and derived more from the field of artificial intelligence).

However such an architecture does impose an extra burden in terms of specifying a lot more knowledge and/or behaviours for the agents for all the relevant contexts. In the simplest case, where one knows what the relevant contexts are and how to recognise them, then the different behaviours can be simply programmed into the agent, along with how to recognise each context, e.g. the approach of the CYC project [13]. Of course, one does not need a specialised architecture to do this – one could just program in more complex rules – but a specialised context-dependent memory might facilitate the process and its checking. In the more complex case, one may not know all the relevant contexts, in which case the agents might need to induce these themselves. This is more complex but possible [6, 8, 19].

Since the behavioural rules at the micro-level of agent-based simulation can be quite specific, I suspect there are quite a few existing agent-based models that have taken some aspects of context-dependency into account without necessarily calling it context-sensitive (e.g. [1, 2, 14]). Each agent in such simulations does detect the kind of situation it is in, and so it will behave differently in different situations. However, these do not distinguish context from other inputs that might influence behaviour and hence do not distinguish what can and cannot be shared between what kinds of situation.

1.4.2 Approaching Context from Qualitative Narratives

One source of information about context that could be exploited to inform the design of social simulations is that of qualitative evidence – that is, the text of observational or ethnographic work or those deriving from the relevant subjects (e.g. in interviews or their online posts). However, such evidence is often treated with suspicion by those who want to be seen to be doing "science". Thus I start this section with a little discussion as to why qualitative research could inform simulation as well as sketching how this might work to a limited degree.

A fundamental value of science is that evidence "trumps" theory, in the sense that if evidence and theory clash then, generally, it is the theory that should be either discarded or modified. One corollary of this is that evidence should not be ignored without a very, *very* good reason. Thus neither qualitative nor quantitative evidence should be ignored. Of course, one should judge the significance of data with respect to its nature and how it was derived, for example, in terms of its relevance, reliability, subject dependence, precision, biases due to observation procedure, distortions in the process of derivation and communication and context-dependency. The quality of data is thus judged in a multitude of ways – different sets of data having different characteristics. Thus qualitative evidence might have a high degree of relevance and precision concerning events that occurred but be subject to different interpretations. Quantitative data is not necessarily more reliable just because it is expressed in a formal, precise form, but it may be if the process by which it is derived is carefully controlled and well founded.

Neo-classical economics has been notorious for ignoring a lot of evidence as to how people make economic decisions. Often this is done via an "as if" argument, which can be roughly expressed as follows: "we know people do not act in this way, but *en masse* we can treat them *as if* they do". In the last couple of decades, experiments have shown that people do not act as if the theory of neo-classical economic decision-making would suggest, e.g. due to different frames of reference [18] or in-group bias [21].

One problem about using qualitative evidence is that it has been difficult to use qualitative data in conjunction with formal modelling methods. However agent-based modelling is well placed to use qualitative evidence to inform the menu of behavioural strategies that people might use in different situations. There is now a growing stream of work on methods to improve the process of the analysis of textual narrative data into behavioural rules (suitable for an agent in an agent-based simulation) – that is, make this more transparent and systematised [5]. Once these behaviours have been incorporated into a simulation at the micro-level, the simulation can be run and then measured to produce numbers that can be compared to macro-level quantitative data [22]. The agent-based simulation can be inspected and experimented upon to understand the process by which this occurs, and the coherence of the qualitative assumptions and the quantitative evidence investigated. Furthermore, a careful analysis of narrative data can suggest some of the context-dependency of behaviour, and if the agents in the model have a context-dependent architecture (as discussed above), this can be incorporated in a *systematic* manner into the model (Fig. 1.3).

Thus explicitly recognising and including context-dependency in formal simulation models can facilitate the integration of qualitative, quantitative and formal modelling approaches in the social sciences. In this way, some of the wealth of qualitative ethnographic, observational and interviewing work that is done in the social sciences can be used to enrich formal simulation models directly, and it does so in a way that allows the quantitative and the qualitative to be assessed together and against each other. The complexity of social phenomena will require all our resources to unpick and understand – facing context-dependency can aid the use of a wider range of evidence without abandoning rigour.

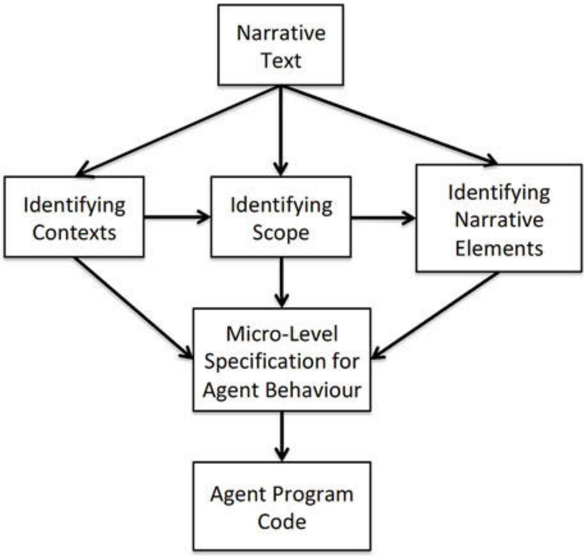


Fig. 1.3 An example (from [5]) of a process of narrative analysis separately identifying context, scope and narrative elements

1.5 Concluding Discussion

Before the advent of cheap computing power, analytic mathematical models were the only formal models available. Solving them or computing answers from them was onerous, so that only simple models were feasible. Their simplicity effectively ruled out context-dependency, leading to a focus on what generic models might tell us. Some of those who appreciated the complexity and context-dependency of social phenomena understandably reacted to this oversimplification and went to the opposite extreme, almost deifying context.

Now that we have cheap computing power, none of this is necessary. We no longer have to distort the phenomena we study in order to achieve useful formal models – we are now free to choose the *most appropriate kind* of formal model – which may well be a computational model such as an agent-based simulation. Cheap computational devices have also resulted in there being a lot more data around about social phenomena – official and informal. We are starting to have enough data to distinguish the different contexts and their associated behaviours – we no longer have to fit generic models to it due to data paucity and limits to the complexity of what we can store/manipulate. Now it is relatively easy to capture, retain, process and compare such data. Finally, we can start to use qualitative and formal methods together – enriching each other. There is no longer any need to ignore context or oversimplify what we observe to obtain and use formal models.

This has been a long time coming, since the old habits derived from a pre-computational age die slowly. However, the age of context-dependent modelling and manipulation is now within our reach. We now longer *have* to avoid it and hope for the best but can start to grapple with its complexity and so make better use of our data (throwing less of it away as noise) and knowledge (bringing more of it to bear down on problems in an more integrated manner). It has the potential for more meaningful, more accurate and more useful models of social phenomena. It will seem odd to future generations that we have been so slow to do this.

Acknowledgements The author acknowledges funding from the EPSRC, grant number EP/H02171X/1, as well as discussion with Emma Norling and a great number of people at the Using and Modelling Context conference series.

References

1. Alam, S.J., Geller, A., Meyer, R., Werth, B.: Modelling contextualized reasoning in complex societies with "endorsements". J. Artif. Soc. Soc. Simul. **13**(4), 6 (2010). http://jasss.soc.surrey.ac.uk/13/4/6.html
2. Antunes, L., Nunes, D., Coelho, H., Balsa, J., Urbano, P.: Context switching versus context permeability in multiple social networks. EPIA. **2009**, 547–559 (2000)
3. Barwise, J., Perry, J.: Situations and Attitudes. MIT Press, Cambridge (1983)
4. Edmonds, B.: Learning and exploiting context in agents. Proceedings of the 1st International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS), Bologna, Italy, July 2002. ACM Press, 1231–1238 (2002)
5. Edmonds, B.: Using qualitative evidence to inform the specification of agent-based models. J. Artif. Soc. Soc. Simul. **18**(1), 18 (2015). http://jasss.soc.surrey.ac.uk/18/1/18.html
6. Edmonds, B., Norling, E.: Integrating learning and inference in multi-agent systems using cognitive context. In: Antunes, L., Takadama, K. (eds.) Multi-Agent-Based Simulation VII, 4442, pp. 142–155 (2007)
7. Fehr, E., Fischbacher, U.: Social norms and human cooperation. Trends Cogn. Sci. **8**(4), 185–190 (2004)
8. Harries, M.B., Sammut, C., Horn, K.: Extracting hidden contexts. Mach. Learn. **32**, 101–112 (1998)
9. Hayes, P.: Contexts in context. Context in Knowledge Representation and Natural Language, AAAI Fall Symposium, November 1997, MIT, Cambridge (1995)
10. Kaelbling, L.P., Littman, M.L., Cassandra, A.R.: Planning and acting in partially observable stochastic domains. Artif. Intell. **101**, 99 (1998)
11. Kim, K.A., Nofsinger, J.R.: The behavior of Japanese individual investors during bull and bear markets. J. Behav. Financ. **8**(3), 138–153 (2007)
12. Kokinov, B., Grinberg, M.: Simulating context effects in problem solving with AMBR. In: Akman, V., Bouquet, P., Thomason, R., Young, R.A. (eds.) Modelling and Using Context, Lecture Notes in Artificial Intelligence 2116, pp. 221–234 (2001)
13. Lenat, D.B.: CYC – a large-scale investment in knowledge infrastructure. Commun. ACM. **38**(11), 33–38 (1995)
14. Nunes, D., Antunes, L., Amblard, F.: Dynamics of relative agreement in multiple social contexts. EPIA. **2013**, 456–467 (2013)
15. Simon, H.A.: Administrative Behaviour, a Study of Decision-Making Processes in Administrative Organization. Macmillan, New York (1947)
16. Tomasello, M.: The Cultural Origins of Human Cognition. Harvard University Press, Cambridge (1999)
17. Turney, P.D.: The identification of context-sensitive features: a formal definition of context for concept learning. In: ICML-96 Workshop on Learning in Context-Sensitive Domains, (Bari, Italy, 1996), pp. 53–59 (1996)

18. Tversky, A., Kahneman, D.: The framing of decisions and the psychology of choice. Science. **211**, 435–458 (1981)
19. Widmer, G.: Tracking context changes through meta-learning. Mach. Learn. **27**, 259–286 (1997)
20. Xenitidou, M., Edmonds, B.: The Complexity of Social Norms. Springer (2014)
21. Yamagishi, T., Mifune, N., Liu, J.H., Pauling, J.: Exchanges of group-based favours: Ingroup bias in the prisoner's dilemma game with minimal groups in Japan and New Zealand. Asian J. Soc. Psychol. **11**, 196–207 (2008)
22. Moss, S., Edmonds, B.: Sociology and simulation: statistical and qualitative cross-validation. Am. J. Sociol. **110**(4), 1095–1131 (2005)

Chapter 2
Agent-Based Modeling with and Without Methodological Individualism

Julie Zahle and Harold Kincaid

Abstract Agent-based models are sometimes associated with a commitment to methodological individualism. In this paper, we explore this linkage through an examination of the following two claims: (1) agent-based models may only be used to provide individualist explanations, and (2) agent-based models should only be employed to offer individualist explanations. We argue that both these claims should be rejected.

Keywords Methodological individualism · Agent-based modeling · Methodological holism · Explanation

2.1 Introduction

Agent-based models (ABMs) are sometimes associated with a commitment to methodological individualism (MI) (see, e.g., [17]:1413; [19]:263; and [21]:188). In this paper, we explore this linkage.[1] We both argue that ABMs need not, and sometimes should not, be seen as embodying the methodological individualist program.

[1]We refer here throughout to ABMs. Most of what we say should be equally applicable to close cousins in computational social science and to parts of evolutionary game theory models, though we do not make that case here.

Julie Zahle and Harold Kincaid contributed equally with all other contributors.

J. Zahle (✉)
Department of Philosophy, University of Bergen, Bergen, Norway

H. Kincaid
School of Economics, University of Cape Town, Cape Town, South Africa

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_2

The discussion that follows is intended for both ABM modelers and philosophers of science. That means at times some discussions will either be obvious or more philosophical than usual for ABM modelers. But we think there is value for both areas in developing bridges.

The MI debate is one of the most long-lasting debates in the philosophy of social science, and we provide an opportunity for ABM modelers to step back a bit and reflect on how their work relates to that debate. Moreover, the debate is not purely intellectual. MI matters to the extent that it encourages ABM modelers to work in certain ways or claim certain virtues for their models. We want to show that their models can meet good scientific standards without fitting MI standards. Even for those ABM modelers not already inclined to be individualist, it is nonetheless still worthwhile to think explicitly about how models capture the social and about the extent to which ABMs need individuals. For social scientists outside ABM circles, we hope to show that ABMs need not be committed to implausible individualist stances and sometimes in fact are not in principle at odds with macrosociological approaches at all.

2.2 Methodological Individualism

Methodological individualism is a claim about explanations. It states that only individualist explanations, that is, explanations which focus on individuals, should be offered in the social sciences. So, for example, an individualist might explain why the government lost the election by claiming that many people were tired of the government's welfare cuts and its tough line on immigrants. Methodological individualists are opposed by methodological holists who contend that not only individualist but also holist explanations should be offered in the social sciences. Holist explanations revolve around social phenomena such as social entities (universities, firms, etc.), social processes (revolutions, migration), and statistical properties (the crime or unemployment rate). If a holist were to explain why the government lost the election, she might point to social factors such as the weak economy, the rise in unemployment, and the opposition of organized labor, thus not mentioning individuals and their reasons to vote as they did. A stronger form of methodological holism states that holist explanations alone should be advanced in the social sciences. This position has few, if any, current proponents, and for this reason, we disregard it for the purposes of this paper.

We have outlined in the foregoing the classic methodological individualism-holism debate. There are also other, more recent versions of the explanatory debate. They concern whether explanations in purely holist terms (i.e., ones that cite causal relations between social entities, etc.) must be supplemented by accounts of the individual-level mechanisms that link the social cause and effect. This is the debate over microfoundations (see, e.g., [11, 12, 23, 24]). The idea of providing microfoundations is a complex one that we cannot delineate here, though we grant the commonsensical point that pursuing the role of individuals as part of social explanations is desirable and fruitful. How ABMs might or might not do so is not our main focus here.

Likewise, we note that there are also individualism-holism debates about theory reduction, ontology, morality, confirmation, heuristics, etc. No doubt there may be interconnections between these different versions and the version we consider, but to pursue a manageable topic, we do not tackle them in this paper. Thus, we restrict our focus to the key MI claim that explanations should be in terms of individuals.

2.3 Agent-Based Modeling

There is hardly any area of social science that has not been influenced by the application of ABMs. Schelling's early models looked at topics that are typically the domain of sociology, and there has been continued work in that field [20]. Nonetheless, there has been as much or more ABM work in all other areas of social science. Demography has been a natural venue for ABMs of course because ABM-like models were already known in the field [3], though those were primarily microsimulations that do not allow interactions between individuals. Political science has seen a boom in the use of agent-based modeling [15], in part because of the pre-existing tradition of modeling political actors with game theory. Mainstream neoclassical economics has long privileged analytic solutions over simulations. But the difficulties confronting equation-based models in incorporating heterogeneity and out-of-equilibrium behavior, among other things, have made ABMs a natural in economics, and there is significant work [7, 8], even if it is still far from the mainstream. Also, there have been many interesting applications in archeology and anthropology [14]), areas whose traditions seem on the whole rather far from the ABM spirit.

The elements of ABMs are widely known to ABM practitioners of course, but not necessarily to other social scientists and philosophers of science. Also whatever the audience, it is useful to clarify, given the looseness with which the ABM terminology can be used. "Model" is in most need of elucidation. The term can be used very widely, but since individualism, as we consider it, is an explanatory thesis, we need to narrow the focus to those things called ABMs that can explain – models no doubt have many other functions. One standard notion of a model is that they are generally abstractions describing entities and relations between them that can be used to potentially explain the world. Models in this sense can run from abstract entities like sets of equations to concrete physical objects like Watson-Crick DNA models. For ABMs, the model objects are the agents and their traits, their environment, and so on. To be more precise, the model is usually the list of agents, initial values of parameters, environments, and rules for agent behavior. Thus to be perhaps pedantically clear, the model is not the running of the simulation nor the reports generated, which are items to be explained, but they are instead the set of statements about agents, environment, etc. at the initial iteration which fix (deterministically or probabilistically) the subsequent realizations of the model.

To further clarify, we should note that the term "agent" has an important but varying and imprecise meaning in different ABMs. The fundamental idea is that the basic entities of the model are "autonomous," though autonomous itself suffers similar vagueness. Yet, there is a core idea: agents are not directed from the top but instead have their own behavior rules. What following their own rules involves can be simple or complex—there may be learning, internal representations of the environment, deductive and other reasoning, and other factors that begin to approximate a self-directed, purposeful human being. ABMs involve different levels and combinations of such internal structure for agents and sometimes have the goal of showing that macrophenomena can arise from relatively unsophisticated agents. But there is still supposed to be a fundamental divide between active agents and a passive environment.

2.4 Are Agent-Based Models Individualistic?

Obviously, ABMs can sometimes provide individualist explanations: the agents in the ABMs may simply be interpreted as individuals and explanations that focus on these individuals may then be offered. The issue we examine in this section is whether ABMs *may only* be interpreted as providing individualist explanations. We argue that this is not the case.

To this end, it is necessary to say something more precise about what to understand by individualist explanations. A quick glance at past and current individualism-holism debates reveals that methodological individualists have endorsed, and been ascribed, many different conceptions of individualist explanations. We want to avoid the confusion that this has often resulted in by distinguishing between several specific versions of individualist explanations. While we don't think that the specifications we go over are equally adopted by ABM modelers, they are notions of individualist explanations that have been widely influential in the general debate.

We proceed by beginning with individualism in its purest forms as it were and then move to more complex views. There is benefit to beginning with the most direct forms of individualism so that we have a clear sense of the logical space of views. Arguments for or against one form of individualism may or may not be relevant to other versions of different logical strength, though later we also give arguments that are generally applicable.

Consider the following version of an individualist explanation:

- Individualist explanations focus exclusively on individuals in isolation, that is, *independently of other individuals and the social and nonsocial physical environment*.

This is individualism as atomism. Less pure forms then allow reference either to other individuals or to the physical environment. Thus, two other individualist theses are:

- Individualist explanations focus exclusively on individuals in *isolation from other individuals and from the social environment*. However, this does not preclude describing the nonsocial physical environment as individuals' context of action.
- Individualist explanations focus exclusively on individuals and their interrelations in *isolation from the social environment and the nonsocial physical environment*.

These are all strict versions of individualism, but they are not purely logical foils. For example, microsimulations that do not allow interaction between individuals embody a form of this strict individualism.

If MI is taken to be one of these theses, then it is relatively obvious that some or most ABMs do not support individualism. The reason is that ABMs may model agents as standing in various relations to each other (e.g., as interacting) and agent's physical circumstances of action (e.g., the vegetation, various tools at their disposal). Explanations based on such models may describe both these aspects, and so they fail to qualify as individualist in the above senses. Since the 1950s at least, many methodological individualists have been happy to refer to relations between individuals in their explanations. For instance, Watkins, a key protagonist of MI in the 1950s, explicitly states that individualist explanations may refer to relations between individuals ([22]:106). In the same vein, there is a long tradition among methodological individualists for permitting individualist explanations to describe individuals' physical surroundings (see [1, 2]). Accordingly, rather few methodological individualists who think that ABMs may only provide individualist explanations will abandon their position in reaction to the observations that ABMs may model individuals' interrelations and/or their physical context of action.

So, contemplate this much broader conception of individualist explanations:

- Individualist explanations focus exclusively on individuals including their interrelations. They may also mention the nonsocial physical environment *as individuals' context of action.*

This version prohibits institutions – an important part of the social environment – from being described as (part of) individuals' context of action. For the present purposes, institutions may be taken to comprise social entities like firms, universities, and states, as well as codified laws and social norms such as traffic regulations and greeting conventions. These are some of the kind of supraindividual factors that individualists typically have tried to avoid in social explanation.

Clearly, ABMs sometimes model institutions in the form of social entities, codified laws, and social norms that may be interpreted as parts of the background against which individuals act. It is perfectly possible for ABM-based explanations to describe institutions and other social phenomena as part of individuals' circumstances of actions, and therefore these explanations disqualify as individualist by the lights of the above definition.

Once more, though, there are MIs who explicitly embrace the view that individualist explanations may actually describe individuals' institutional context of action. This is compatible with these explanations focusing on individuals, these MIs think, because the institutions only set the stage for individuals' action: they are relegated to a secondary role as part of the circumstances of individuals' action and are passive parts of the environment that are used and constrain individuals' actions in the same way as nonsocial parts of the environment. The position is commonly referred to as institutional individualism. It was first sketched by Popper (or even perhaps Weber) and subsequently elaborated by Agassi and Jarvie [1, 2, 10]. Evidently, nothing in the foregoing will convince institutional MIs to give up on the claim that ABMs may only be interpreted as providing individualist explanations. Every explanation takes something as unexplained, so there is no inherent problem for institutional individualists in leaving some social things unexplained that are used as the background against which individuals act.

In light of these considerations, it is natural to wonder whether there are any reasons against the claim that ABMs may only be used to provide individualist explanations that apply to all individualists, including the most inclusive formulations of the position. We think there are several.

First, and obvious to some but not all agent-based modelers, the agents in ABMs need not be interpreted as individual agents. All the agents may be collective agents, that is, social entities of some sort. Here collections of unnamed individuals and their relations are both left hidden in a black box, underneath descriptions of the social entities they constitute. Thus, for example, all of the following have been agents in ABMs in the literature: households, firms, nations, banks, central banks, political parties, elites, and governments [4, 8, 15]. In this manner, ABM-based explanations may focus exclusively on social entities and their properties (as opposed to leaving them in the unexplained background), and for all MIs, this is unacceptable: these explanations are straightforward holist explanations focusing on social phenomena in the form of social entities and their properties.

Moreover, ABMs may be mixed in their ontology, allowing both individual agents and social entities that are agents as well. If they are interpreted as containing such collective agents too, ABM models are not compatible even with the weakest forms of individualism, namely, institutional individualism, for the latter takes social structure to be a passive background against which agents act much like they might act against a background of resources in the environment. If an economic model, for example, includes both individual investors and banks, governments, etc., all of which have their own rules for interacting with each other, then it does not explain purely in individualist terms.

A second, general way that ABMs may not be individualist is that sets or collectives of individual agents may be ascribed various aggregative properties. An aggregative property, as we use the term here, is one that is expressed in terms of individuals and operations on them to produce traits, such as the unemployment level, and is to be distinguished from directly positing traits of social entities such as democratic versus authoritarian states where there is no direct mention of individuals. Most notably, these aggregate properties include statistical properties (there is a certain rate of unemployment among the agents, say) and formal network properties (e.g., the relations among individuals constitute a network that has a certain density). According to all MIs, any aggregate properties of sets or collections of individuals are unacceptable if they somehow are part of the focus of the explanation (i.e., not backgrounded in explanation). Obviously, it may be difficult to draw the distinction between what is foregrounded and backgrounded in explanation – a distinction institutional MIs insist on – but we do not need to address this issue here, since it is one drawn by the individualist. What matters is that if we grant the distinction, then if these properties of sets or collections of individual agents figure as part of the focus of the explanation, they are unacceptable even to institutional individualists.

2.5 Should Agent-Based Models Be Individualistic?

The foregoing discussion has shown that ABMs may be interpreted as providing not only individualist but also holist explanations. This still leaves it open to MIs to argue that ABMs *should only* be used to offer individualist explanations because holist accounts violate general standards of good explanation.[2]

One way in which MIs might push this point is by first noting that, like individualist explanations, holist explanations are causal explanations: they state that social phenomena causally contribute to the bringing about of events. MIs may then continue that this makes these explanations problematic because social phenomena are causally inert. Individuals alone are causally effective, and for this reason, it is only legitimate to use ABMs for the purposes of offering individualist (causal) explanations.

In order to motivate this contention, MIs may claim that social entities, and sets and collections of individuals, have their properties in virtue of individuals' having various properties. Social properties somehow noncausally (synchronically) depend on individualist properties (i.e., individuals' properties). Commonly, this idea is further cashed out in terms of supervenience. Accordingly, social properties are said to supervene on individualist properties meaning, roughly, that there can be no change at the level of social properties unless there is also a change at the level of individualist properties. The point is now that MIs may maintain that because social properties supervene on individuals' properties, they are causally inert: all the causal work is done by individuals.

For the present purposes, there is no need to go into the debate about whether supervenient properties are indeed causally inert. It suffices to note that MIs would be ill-advised to go down this road. The reason is that just as social properties supervene on individualist properties, the latter supervene on biological properties that, in turn, supervene on chemical properties that supervene on physical properties. Or, so proponents of supervenience often argue (see, e.g., [9]). As a result, individualist properties, qua being supervenient, are causally inert too, and MIs must admit that it isn't legitimate to offer individualist (causal) explanations either. Note

[2]The following discussion is an extension of earlier work in [24].

that these considerations potentially generalize. MIs may propose another way in which to spell out the idea that social phenomena lack causal efficacy because they noncausally (synchronically) depend on individuals. In that case, however, MIs must make sure that their point about causal inertness does not extend to individuals so that these are robbed of their causal efficacy as well. It is doubtful that MIs may successfully meet this challenge (see, e.g., [6]).

MIs' conviction that social phenomena are causally inert may be further challenged by registering that it is not supported by standard conceptions of causation. To see this, consider the regularity account, which states, roughly, that x causes y if x is regularly followed by y. By the lights of this account, for example, "an increase in the money supply causes increases in economic output," makes a causal claim. Also, it is reasonable to hold that manipulating the money supply by a central bank would result in an increase in economic output. Thus, x – increase money supply in the example – would hereby count as the cause of y – increase in economic output – following the manipulationist account of causation. Similar points may be made by adopting yet other conceptions of causation such as the mechanism account (which says roughly that x is a cause of y if there is a mechanism going from x to y) and the counterfactual view (which states roughly that x causes y if it is true that had x not occurred, then y had not occurred either). The upshot of these reflections is that, as it stands, there are no convincing reasons for holding that ABM-based holist (causal) explanations should not be advanced.

A second way in which to defend the view that ABMs should only be used to offer individualist explanations might be to argue that holist explanations treat social entities as intentional agents, that is, as entities that have a mental life in the form of beliefs, desires, and rationality. According to MIs, collective agents do not have these elements and thus cannot be agents in ABMs. That is, we have already argued that ABMs with collective entities can provide legitimate causal claims. The formalism of ABMs does not tell us how we are to interpret the entities that stand in causal relations. Individualists would like to claim that it is illegitimate for ABMs to have basic actors that are not intentional agents in the sense described a second ago.

This contention can be met in three ways:

1. By arguing that collective entities can have full-bodied beliefs, desires, and rationality as traditionally understood despite appearances to the contrary
2. By defending the naturalist philosophical position that notions of intentional agency are what our best empirical cognitive science and psychology finds most explanatory, and from that perspective, there are legitimate explanatory ways to think of collective social entities as intentional agents
3. By maintaining that in fact, and perhaps unrecognized by individualists, existing successful ABMs themselves do not always treat their individual human agents as fully intentional – they do not give an account that makes their basic actors fully rational, etc. – and thus ABMs cannot require that its basic actors be intentional agents in the strong sense individualists need. Actually, a major

motivation of ABMs is to show that macrophenomena can emerge from agents that are not nearly as rational, sophisticated intentional agent as pictured, for example, by neoclassical economics.

The first line of argument has been developed by List and Pettit [16]. They contend that an intentional agent – an agent who consciously acts purposefully – is distinguished by having representational and motivational states and by having the capacity to process and act on these states. Further, they maintain, groups with a certain form of organization meet these conditions of intentional agency. List and Pettit's is a conceptual analysis project. We are much more sympathetic to the naturalist approach.

The naturalist approach asks what we gain by positing the characteristics associated with agents such as beliefs, desires, plans, and so on. A key question is whether doing so allows us to find usable patterns in the behavior of the entities we are describing. Dennett, who is a prime expositor of this view, spells out "usable patterns" in terms of real patterns, which are given an explication in terms of nonredundant information [5]. Revealed preference theory in economics, which tries to give explanations entirely in terms of observable individual choices, provides a particularly clear formulation: if a set of choices by an entity meet certain conditions, then there is proof those choices can be represented by a utility function which will describe any entity with that set of preferences. However, real patterns and utility functions in principle are not restricted to individual human agents. Thus, collective social entities in principle can be agents if empirically there are real patterns, for instance, utility functions that describe their behavior. Work on ABMs across social science disciplines shows that there may well be the empirical evidence that is needed.

Finally, ABMs and those who defend them as individualist do not in practice apply these requirements to the agents of the models, even when they are to be interpreted as individual human agents. Many agents in ABMs are what Salamon [18] calls reactive agents that do not even approximate the sophisticated mental life of real human actors. Moreover, even the most sophisticated of ABM agents ("deliberative agents" in Salamon's terms) are very far from having the complex desires, beliefs, and reasoning that are thought essential to the philosophical concept of an intentional actor.

Of course, there may be still further general constraints on explanation that defenders of MI might try to use to show that apparently holist (pure or partial) ABMs are illegitimate. For instance, there is a debate on mechanisms and explanatory completeness which might support the individualist, but we have argued elsewhere that they are unclear and unhelpful to the individualist [13]. We will not try to run through all such possible relatively a priori ideas about explanation in the space allowed, but we note that the naturalist presumption is against philosophers ruling that apparently successful science is inadequate on philosophical, extrascientific grounds.

2.6 Conclusion

In this paper, we have examined the relationship between ABM and MI. First, we examined the claim that ABMs *may solely* be used for providing individualist explanations. We argued that ABMs may equally well be employed to offer holist explanations. In this process, we showed that some reasons against the claim that ABMs may only serve as basis for individualist explanations have a limited scope (they only apply if certain versions of individualist explanations are adopted), whereas other reasons count against all forms of MI (all versions of individualist explanations). Second, we discussed the claim that ABMs *should only* be used for offering individualist explanations. Here, we argued that two significant individualist arguments (by appeal to causation and intentional agency, respectively) should be rejected. Evidently, there is much more to be said on the relationship between ABM and MI. Still, we hope to have made a good start on reflecting on the extent to which ABMs may, and should, be employed to offer explanations that focus on individuals and social phenomena, respectively.

References

1. Agassi, J.: Methodological individualism. Br. J. Sociol. **11**(3), 244–270 (1960)
2. Agassi, J.: Institutional individualism. Br. J. Sociol. **26**(2), 144–155 (1975)
3. Billari, F., Praskawetz, A. (eds.): Agent-Based Computational Demography. Springer, New York (2003)
4. Chen, S.-H.: Agent-Based Computational Economics: How the Idea Originated and Where it Is Going. (Routledge Advances in Experimental and Computable Economics Book). Routledge, New York (2017)
5. Dennett, D.: Real patterns. J. Philos. **88**, 27–51 (1991)
6. Elder-Vass, D.: The Causal Powers of Social Structures. Cambridge University Press, Cambridge (2010)
7. Gallegati, M., Palestrini, A., Russo, A.: Introduction to Agent-Based Economics. Academic Press, New York (2017)
8. Hamill, L., Gilbert, N.: Agent-Based Modelling in Economics. John Wiley, New York (2016)
9. Jackson, F., Pettit, P.: In defense of explanatory ecumenism. Econ. Philos. **8**, 1–21 (1992)
10. Jarvie, I.C.: Concepts and Society. Routledge and Kegan Paul, London (1972)
11. Kincaid, H.: Philosophical Foundations of the Social Sciences: Analyzing Controversies in Social Research. Cambridge University Press, Cambridge (1996)
12. Kincaid, H.: Individualism and the Unity of Science. Rowman and Littlefield, Lanham (1997)
13. Kincaid, H.: Open empirical and methodological issues in the individualism-holism debate. Philos. Sci. **82**(5), 1127–1138 (2015)
14. Kohler, T., Gumerman, G. (eds.): Dynamics in Human and Primate Societies: Agent Based Modeling of Social and Spatial Processes. Oxford University Press, New York (2000)
15. Laver, M., Sergetti, E.: Party Competition: An Agent-Based Model. Princeton University Press, Princeton (2012)
16. List, C., Petit, P.: Group Agency. Oxford University Press, New York (2011)
17. O'Sullivan, D., Haklay, M.: Agent-based models and individualism: is the world agent-based? Environ Plan A. **32**, 1409–1425 (2000)

18. Salamon, T.: Agent-Based Models: Developing Computer Simulations for a Better Understanding of Social Processes. Academic Series, Zvonin (2011)
19. Sawyer, R.K.: The mechanisms of emergence. Philos. Soc. Sci. **34**(2), 260–282 (2004)
20. Squazzoni, F.: Agent-Based Computational Sociology. Wiley, New Work (2012)
21. Wan, P.Y.: Emergence à la systems theory: epistemological Totalausschluss or ontological novelty? Philos. Soc. Sci. **41**(2), 178–210 (2011)
22. Watkins, J.W.N.: Historical explanations in the social sciences. Br. J. Philos. Sci. **8**(30), 104–117 (1957)
23. Zahle, J.: Methodological holism in the social sciences. In: Zalta, E.N. (ed.) Stanford Encyclopedia of Philosophy, http://plato.stanford.edu/, pp. 1–39 (2016)
24. Zahle, J., Kincaid, H.: Why be a methodological individualist? Synthese. **196**(2), 655–675 (2019)

Chapter 3
Inflation Expectations in a Small Open Economy

Emiliano Alvarez, Juan Gabriel Brida, and Marco Dueñas

Abstract This paper studies the dynamics of the formation of inflation expectations in small and open economies, based on an agent-based model (ABM). Agents are assumed to have a limited rationality to adapt their strategies over time and to have difficulty at perceiving the signals from the Central Bank (CB). The CB tries to achieve a target inflation by setting an inflation rate. The monetary authority is committed to a certain level toward which it is directed with monetary policy instruments – the interest rate, in this economy. To achieve this goal, it is essential that the expectations of the private sector get aligned with the objective proposed by the CB. The CB uses a Taylor rule for open economies in order to determine the interest rate.

In the case of small and open economies, the inflationary pressures come from internal and external factors. We use a *New Keynesian* (NK) model with nominal rigidity and adaptive agents in order to study the alternatives of the CB in order to achieve its objectives. Our results show that a high level of credibility of the monetary authority and stable external conditions are necessary for the success of the policies promoted by the CB.

Keywords Agent-based model · Inflation expectations · Social preference · Economics

E. Alvarez (✉) · J. G. Brida
GIDE-DMMC, Facultad de Ciencias Económicas y de Administración, Universidad de la República, Montevideo, Uruguay
e-mail: ealvarez@ccee.edu.uy; gbrida@ccee.edu.uy

M. Dueñas
Department of Economics and Trade, Universidad de Bogotá Jorge Tadeo Lozano, Bogotá, Colombia

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_3

3.1 Introduction

In the last decades, monetary policies have based their strategies to control inflation by the management of expectations [12]. A communication strategy used by CBs has been through an explicit target of inflation, which serves as a reference to the private sector. In this model, the CB announces the inflation target and uses the interest rate as a policy instrument. It uses a Taylor rule for open economies, introduced in the literature by [2, 10]; and [11].

The economy is modeled as an adaptive system, where individuals and firms follow behavioral heuristics under constant learning. This framework allows us to study the interaction between individuals and the aggregate behavior resulting from this interaction. Starting from an NK model [12], and following [8] and [9], which extended the results to a small and open economy, we studied the behavior of individuals and the degree of adjustment of the inflation to the objective proposed by the CB.

Households consume goods that might be produced at home or in a foreign country. The households' reservation wage for the next period is based on expected inflation; however they do not perceive exactly the CB signal [1, 8, 9]. More precisely, individuals do not have the certainty that the CB will achieve the objective, and, for this reason, they base their expectations on both the objective and the past inflation [8, 9]. Therefore, the reservation wage is defined according to

$$w_{i,t} = w_{i,t-1}(1 + \gamma^k_{i,t}\pi^e_{i,t+1}), \tag{3.1}$$

where $\pi^e_{i,t+1}$ is the expected inflation. This results in heterogeneous reservation wages from heterogeneous expectations, which is corroborated from an empirical point of view from surveys [5, 7] and experiments [6].

Firms have a cost function that depends on the amounts and prices of physical and human capital, using a markup μ on their average costs to choose the price of the good. It is assumed that home does not influence the price of the imported good, since it is a small economy.

The households' nominal income comes from different sources: their salary ($w_{i,t}h_{i,t}$) when hired, the share of firms' income in the previous period (Π_{t-1}/N), and savings (b_i).

The consumption and savings decisions of individuals depend on their permanent income, which is a weighted average of the past nominal income, adjusted by inflation. Individuals decide to consume a b proportion of their permanent income, which depends on the differential between the interest rate and the expected inflation.

Individuals obtain a greater utility through the consumption of locally produced goods. The strategies of the individuals (in terms of the choice of their reserve salary and the proportion of permanent income dedicated to the consumption of the good) are key for this objective. In this economy, successful strategies are more likely to be imitated than unsuccessful ones. Consumption, wages, and, as a consequence, prices are the result of the process of interaction of individuals.

Each of the decision-making agents follows certain heuristics:

- The Central Bank sets the interest rate i_t using a Taylor rule for open economies, with the monetary policy parameters ϕ_π, ϕ_u, and ϕ_e. It can be written as:

$$1 + i_t = (1 + \pi^T)(1 + r_t^n) \left(\frac{1 + \pi_t}{1 + \pi^T} \right)^{\phi_\pi} \left(\frac{1 + u^*}{1 + u_t} \right)^{\phi_u} \left(\frac{1 + e_t}{1 + e_{t-1}} \right)^{\phi_e} \tag{3.2}$$

  where π^T is the inflation target, r^n the natural level, u^* the natural rate of unemployment, and e the real exchange rate.
- Households learn from the successful strategies of others, through imitation. The nominal rigidity and the inertia of past inflation in wages are a consequence of the heuristic used to choose the *reserve* wage.
- Firms increase or decrease their physical and human capital to the following period according to the relationship between the benefits of the previous period and the weighted average adjusted for inflation.

Then, it is important to establish if Central Bank credibility is endogenous to its performance. As [3] states, the fulfillment of the goal in the past is important to credibility. Also, we should expect systemic effects: a phase transition between confidence and confidence loss.

Central Bank's credibility (χ_{it}, Eq. (3.3)) determines whether agents in the next period will be based mostly on the target or trend inflation (Eq. (3.4)). It is measured each period. We assume $\chi_{\text{it}}^{\text{min}} = 0.1$ and $\chi_{\text{it}}^{\text{max}} = 0.9$, with Δ_i as tolerance to the Central Bank.

$$\chi_{\text{it}} = 1 - \frac{|\pi_t^T - \pi_t|}{\Delta_i}, \text{ with } \Delta_i \sim U[0, \delta] \tag{3.3}$$

Note here that unlike [8], credibility is time-dependent; this implies that Central Bank's credibility arises from their performance. There is a difference with other approaches, where credibility is generated from Central Bank's signals and not from the results of its policy.

Parameter Δ_i, defined as individual tolerance to the Central Bank policy, is an indirect measure of the maximum deviation from the target so that the target is no longer credible. It is expected that this value will depend on different factors intrinsic to the individuals, as well as will be different according to the countries and the tradition of fulfillment of the goals of Central Banks.

$$\pi_{i,t+1}^e = \pi_{\text{it}}^T \chi_{\text{it}} + \pi_{\text{it}}^{\text{trend}} (1 - \chi_{\text{it}}) \tag{3.4}$$

$$\pi_t^{\text{trend}} = (1 - \rho)\pi_{t-1} + \rho \pi_{t-1}^{\text{trend}} \tag{3.5}$$

Expected inflation for the next period (Eq. (3.4)) is a weighted average of past inflation and the monetary policy objective. Unlike [4] – such as $0.1 < \chi_{\mathrm{it}} < 0.9$- in the face of high inflation, expected inflation will not be above trend. Trend inflation (Eq. (3.5)) is a weighted average of the time series of inflation. It is understood that individuals not only look at the latest data – which may be subject to short-term variations – but take into account the previous behavior of the variable. It plays a fundamental role ρ, since it is the parameter of temporal preference: a value of ρ close to one shows little variability of the trend with respect to short-term variations; a ρ close to zero, on the other hand, makes the trend very sensitive to short-term shocks. As shocks in trend inflation cause changes in expectations, it is important to take into account the results of the model when facing changes in ρ.

3.2 Results

3.2.1 Stability Under Households Decision

First of all, we want to know if the system is stable or prone to hyperinflationary crises based on the structure and value of the parameters of the model. In particular, we will study the mechanisms by which inflation expectations differ significantly from objectives and for which values of individual parameters monetary policies fail to control the phenomenon. We emphasize the parameters ρ and δ since they are intrinsic to individuals, i.e., there are no policies that modify their values in the short term. For this reason, policies can be observed that in some cases can give excellent results, while in others they do not reach the objectives. The ρ parameter affects expected inflation from the link between trend inflation and present inflation, while the ρ parameter affects expected inflation from the credibility of monetary policy.

In Fig. 3.1 is observed the proportion of simulations with stable – non-hiperinflationary – outputs. What we can notice is that the system stabilizes when there is more tolerance of households to Central Bank's deviations in the control of inflation (greater δ) and when households take more into account trend inflation than present inflation (greater ρ). Both results show that short-term fluctuations in inflation can generate crises in the system. More importantly, the same shock does not affect the system if households have a high tolerance to the Central Bank or if current inflation is of little importance compared to trend inflation.

As each of these parameters is positively correlated with the stability of the model, their joint analysis allows establishing zones in the $(rho, delta)$ space where the system is stable. The aggregate analysis of these variables is much clearer in the equation of expected inflation. With ρ close to zero, $\pi_{\mathrm{it}}^{\mathrm{trend}}$ depends more on short-term shocks. On the other hand, a lower δ (global tolerance) implies lower Δ_i (individual tolerance), because of the way the model is defined. *Ceteris paribus*, a lower δ, generates lower credibility of the Central Bank. Finally, a

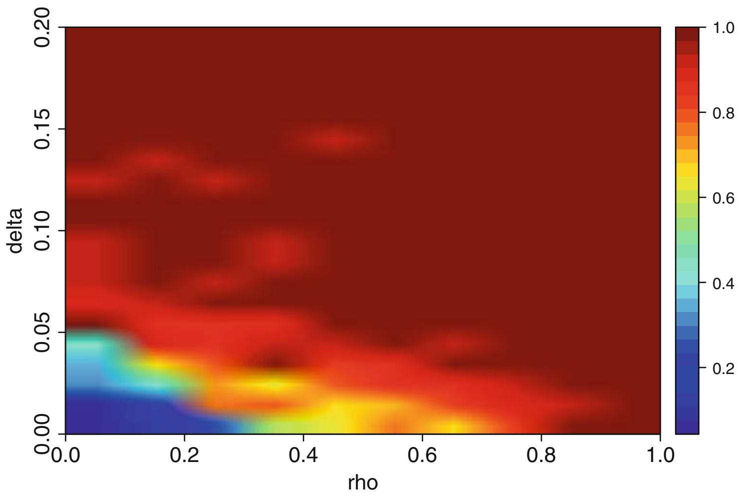


Fig. 3.1 Proportion of stable output, according to δ and ρ

lower credibility of the CB together with a greater prevalence of short-term shocks generates conditions of more instability in the economy.

3.3 Concluding Remarks

Therefore, this paper proposes a bottom-up approach to study the heterogeneity of inflation expectations in open economies, with implications for the policies to be adopted.

In the first place, differences in expectations formed by agents are not only explained by differences in information or lags at the moment of making the decision but from an evolution of the strategies of agents and the implicit heterogeneity in them. Even with perfect information, the differences in the credibility of the Central Bank generate consequences in expectations.

Differences in the exposure of workers to prices cause distortions in the strategies that maximize their utility. This generates changes in consumption and savings decisions throughout the economy, with effects that are amplified and can make a monetary rule à la Taylor unfeasible. In economies subject to price shocks, it is important to consider this source of instability when establishing stabilization plans.

Another aspect we discuss in this work is the relationship between the speed of the processes that occur in the labor market and the variations in internal prices. A labor market that is very sensitive to variations in corporate earnings causes greater fluctuations in demand and prices. Finally, it affects the agents' expectations and harms the stabilization carried out by the Central Bank.

Preliminary results show that the system becomes unstable with less tolerant households, also with households more sensitive to short-term fluctuations. Estimation errors, their heterogeneity between households, and Central Bank's credibility are linked to δ. At the same time, there is a dilemma for Central Banks: decide an acceptable range of their policies, which can be an anchor for inflation expectations and also an anchor for individuals' tolerance to deviations.

One possible extension consists of studying this decision model empirically. Several papers have shown that the hypothesis of rational expectations with complete information is rejected. In the economic literature, this issue is explained as an asymmetric information problem. A research line to follow is to analyze whether the bias in expectations is related to characteristics intrinsic to individuals rather than to publicly available information.

References

1. Agliari, A., Massaro, D., Pecora, N., Spelta, A.: Inflation targeting, recursive inattentiveness, and heterogeneous beliefs. J. Money Credit Bank. **49**(7), 1587–1619 (2017)
2. Ball, L.M.: Policy rules for open economies. In: Monetary Policy Rules, pp. 127–156. University of Chicago, Chicago (1999)
3. Blinder, A.S.: Central-bank credibility: why do we care? How do we build it? Am. Econ. Rev. **90**(5), 1421–1431 (2000)
4. Bouchaud, J.-P., Gualdi, S., Tarzia, M., Zamponi, F.: Optimal inflation target: insights from an agent-based model. Economics: The Open-Access, Open-Assessment E-Journal. **12**(2018–15), 1–26 (2018). https://doi.org/10.5018/economics-ejournal.ja.2018-15
5. Carroll, C.D.: Macroeconomic expectations of households and professional forecasters. Q. J. Econ. **118**(1), 269–298 (2003)
6. Hommes, C.: The heterogeneous expectations hypothesis: some evidence from the lab. J. Econ. Dyn. Control **35**(1), 1–24 (2011)
7. Mankiw, N.G., Reis, R., Wolfers, J.: Disagreement about inflation expectations. NBER Macroecon. Ann. **18**, 209–248 (2003)
8. Salle, I., Yıldızoğlu, M., Sénégas, M-A.: Inflation targeting in a learning economy: an ABM perspective. Econ. Model. **34**, 114–128 (2013)
9. Salle, I.L.: Modeling expectations in agent-based models: an application to Central Bank's communication and monetary policy. Econ. Model. **46**, 130–141 (2015)
10. Svensson, L.E.: Open-economy inflation targeting. J. Int. Econ. **50**(1), 155–183 (2000)
11. Taylor, J.B.: The role of the exchange rate in monetary-policy rules. Am. Econ. Rev. **91**(2), 263–267 (2001)
12. Woodford, M.: Interest and prices: foundations of a theory of monetary policy. Princeton University Press, Princeton (2003)

Chapter 4
Qualitative Data in the Service of Model Building: The Case of Structural Shirking

Patrycja Antosz and Harko Verhagen

Abstract This chapter shows how qualitative data can inform building computational models. The general issue is illustrated with the example of a model of structural shirking in organisations, i.e. insufficient time and effort stemming from the structure of the performed work. The first attempt to build a model of shirking with the use of assumptions present in social scientific theories displayed many shortcomings. Thus, a mixed-methods approach was chosen to inform the development of a second computational model. Conceptualising the second model began with performing individual IDIs with managers and lower-level employees and augmenting them with analyses of Polish legislation regulating employment relationships. Initial findings were enriched with theoretical assumptions. The complete concept of the mechanism of structural shirking was operationalised as a computational model. Having developed both models, we discuss the phenomenon of shirking informing theories and real-world practices, as well as ways to study these practices in novel forms.

Keywords Shirking · Work performance · Mixed methods · Empirical-based model · Theory-based model

4.1 Introduction

This is a story of a failure. Of developing a model based on theories providing sound, yet invalid assumptions. Of having to take a step back and reconceptualise the entire study. A very difficult decision, but one that needed to be made. Of how

P. Antosz (✉)
Department of Social Sciences, University College Groningen, Groningen, The Netherlands
e-mail: p.p.antosz@rug.nl

H. Verhagen
Department of Computer and Systems Sciences, Stockholm University, Kista, Sweden
e-mail: verhagen@dsv.su.se

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_4

qualitative methods can be useful in such times of despair. Finally, of how a failure can be transformed into a success.

Shirking is the main character in the story, the studied phenomenon. It occurs when employees voluntarily work less than agreed upon. Shirking has been a subject of scientific inquiry of several disciplines, including sociology of work, sociology of organisations, management, social psychology, and economics. Given the wealth of theories originating from different perspectives and assumptions, it seems that there would be a good chance of conceptualising a theoretical model explaining the phenomenon. Yet, more is not always better. Plurality generated a lack of agreement about the definition of the concept, as well as the mechanisms causing shirking. Status quo creates an opportunity to show the strengths of applying computational modelling to solving unsettled problems of definition in social sciences. It allows for presenting a clear operationalisation of, in our case, shirking, identifying valid indicators, estimating the magnitude of the phenomenon and tracing its dynamics.

The tale told here does not, however, focus on the main character. Shirking serves merely as an example of how theoretic progress can be made with the use of a mixed-methods approach. The chapter is organised as follows: in Sect. 4.2 we present some of the theories mentioned above in more detail. Following this, we describe the findings of a simulation study based solely on assumptions present in these theories. The shortcomings of the model inspired turning to the empirical world (Sect. 4.4) to take a closer look at reality in order to address a redefinition of the concept of shirking and discover the mechanisms involved in it. Reconceptualisation began with analyses of a legal perspective, and of perspectives of managers and employees. A second model, described with the results of the simulations in Sect. 4.5, brought together the findings of the empirical study and assumptions from relevant theoretical perspectives. The story concludes discussing the findings of both simulation studies and the consequences for the existing theories. It also summarises how the process informed a new take on measuring and studying the phenomenon "in the wild" and, finally, potential consequences for the real world.

4.2 The Phenomenon Under Investigation: Shirking

The first obstacle in furthering knowledge on shirking stems from a lack of an agreed-upon definition of the phenomenon. Different authors departing from a range of theories and disciplines have offered various propositions. The most popular conceptualisations focused on *effort* [1–3]. Shirking is therefore usually considered a tendency for workers to give less than a full, or agreed upon, level of effort. Recently, another school of thought on shirking emerged, where the emphasis shifts from effort to *time*. Roland Paulsen [4] introduced the term empty labour, i.e. everything you do at work that is not your work [4, p., 5]. An employee starts shirking as soon as he or she starts doing something other than work tasks. So far, no effort to consolidate the two approaches, focused on effort and on time, has taken place.

The second obstacle is related to a lack of a coherent concept of the mechanisms leading to shirking behaviour. Proposed models of shirking concentrated mainly on discovering singular causal relationships. As a result, various determinants of task performance have been identified. They include monitoring practices [5], level of unemployment [6, 7], payment scheme [8], relationship between the manager and the employee [9], expected level of effort [10], individual differences among employees [11], and organisational structure [12].

Developing a thorough computational model of shirking in organisations has to coherently combine assumptions present in these numerous lines of research. Analysing several approaches in a multidisciplinary fashion assumes that more scientific knowledge is already present, but only by putting it in a specific context can we draw relevant conclusions regarding shirking.

4.3 Starting from Theories: Simulation Model of Shirking #1

In the initial model, shirking is interpreted as a portion of the employee's time devoted to shirking activities, i.e. non-work-related activities. The remaining part of time is devoted to working. In an organisation comprising100 workers randomly allocated on a 25 × 25 torus, employees are characterised by individual probabilities of shirking. Initial probability of shirking is ascribed to employees according to a Beta distribution. Employees make decisions regarding shirking on the basis of peer influence – employees learn what shirking levels their close colleagues find acceptable and imitate each other's behaviours [13]. Influence of information operates within a 2.0 radius and decreases exponentially with distance from a given colleague. Employees do now know the expectations of their manager regarding the amount of working time [14]. If organisational productivity, realised as a sum of employees' time devoted to working, is lower than the manager's expectations, a certain proportion of employees is monitored. Monitoring probability is dynamic and its initial value reflects individual differences among managers and original trust in the selected team of subordinates. Employees found not performing up to the manager's standards are selected for dismissal. Before being terminated, employees warn their close colleagues to increase working. Initial shirking level of new employees who replaced the identified shirkers reflects the conditions on the local labour market [7]. The number of monitored employees increases if productivity has been decreasing for the past three time steps or if it fell drastically (by more than 10%) in comparison to the previous realisation. The simulation stops when a stable equilibrium between employer's expectations regarding productivity and shirking among employees is reached, i.e. when (a) monitoring stops, indicating full trust in employees' performance and (b) employees cease to change their shirking levels.

4.4 Findings of Model #1 and Shortcomings

Initial simulations of the model confirmed previous findings. Convergence between the manager's expectations and the employees' actions occurs faster and generates smaller costs in case of lower discrepancies between initial respective values. Also, increasing monitoring and replacing shirking employees with hard-working ones accelerate reaching the equilibrium. It became clear that the results of the simulations were trivial and did not offer much additional insights into the practices of shirking. Some findings were counterintuitive. For example, in scenarios when it took longer to achieve stability, 10,000 employees were terminated in an organisation of 100 workers. An assumption that shirking leads to dismissal, present in several economic theories, e.g. inspection games [15] and efficiency wages [7], was not realistic in the context of this model, thus further development was halted, and reconceptualisation of the research problem took place.

4.5 Goal of the Study

The futility of the first model left us stranded. It became clear that a complete reconceptualisation of the research problem is needed. Gaps in the existing scientific knowledge were broad and included both details of existing conceptualisations and issues related to the mechanism through which shirking in organisations occurs:

1. The concept of shirking
   (a) Is there a legal basis for conceptualising qualitative and quantitative dimensions of shirking?
   (b) Are contractual obligations the reference points for shirking activities?
2. The mechanisms involved in shirking
   (a) Through what process does shirking arise?
   (b) What are the most important determinants of shirking?
   (c) What decision rules guide actions of employees and managers?

In order to address the gaps in existing knowledge and conceptualise the new model of shirking, two types of qualitative studies were performed: analyses of Polish legislation governing employment relationships and a series of individual in-depth interviews with two groups, i.e. lower-level employees, employed on a basis of a work contract, and managers, supervising the work of at least three employees, who are employed on a basis of a work contract and perform similar tasks.

The second stage of the study involved using the knowledge acquired in the qualitative phase as a context and supplementing it with assumptions existing in social scientific theories in a more productive manner, in order to specify the mechanism of shirking, obtain more reliable estimates of the phenomenon and trace its dynamics.

4.6 Turning to the Empirical World: Developing Model #2

4.6.1 Addressing Gap 1: Legislation Analysis to Define the Concept of Shirking

Shirking constitutes a dimension of organisational culture, which is realised on three levels: the legislative level, the managerial level, and the employee level. The legislative system defines norms that should uniformly be implemented in labour market practices. These norms govern contracts between employees and their employer in organisations. Managerial expectations, regarding the fashion in which daily tasks are to be completed, are located on a level below the official contractual agreement. On the lowest level of the normative hierarchy, employee actions take place. The next section determines whether theoretically identified dimensions of shirking reflect provisions present in the legislation in Poland, where our case study took place. Moreover, it resolves the issue of the reference point for diagnosing shirking behaviour through verifying what should employee actions be contrasted with – the contractual agreement or the manager's expectations.

Description of the normative layer regulating specific characteristics of the employment relationship[1] related to shirking serves as a general summary and is narrowed down to employment relationships as regulated by the Labour Code[2] (the Labour Code or the L.C., hereafter). By establishing an employment relationship, the employee agrees to perform work of a specified type for the benefit of the employer and under his supervision, in a place and the times specified by the employer; the employer undertakes to employ the employee in return for remuneration (art. 22 §1 of the L.C.). Basic obligations of employees are enumerated in the Labour Code, which requires that they perform work conscientiously and carefully, and comply with the instructions of superiors related to the work, as long as the instructions are not contrary to the provisions of law or the employment contract (art. 100 §1 of the L.C.). Qualitative dimension of working and shirking is therefore specified through the concepts of conscientiousness and carefulness. The former designates a proper work attitude and considers individual characteristics of the employee, such as work experience, psychophysical abilities, qualifications, or age [17]. The latter is related to the rules of knowledge and experience that point to an objectively proper work conduct. Those rules are created socially, particularly in labour courts. In the eyes of the Supreme Court of the Republic of Poland, careful work is characterised by adhering to technical (praxeological) rules and rationality [17]. The employee cannot take responsibility for inadequate

[1] Other contractual forms, such as project contract or a contract of mandate, are regulated by the Civil Code and do not evoke an employment relationship and therefore are excluded from the presented analyses.

[2] Act of 26 June 1974 on the Labour Code (J.L. 1974 No. 24, item 141), consolidated text of 4 November 2014 (J.L. 2014 No. 0, item 1502). Translations of the Labour Code after Agnieszka Jamroży [16].

carefulness so long as he or she is performing work conscientiously. Therefore, an obligation of conscientiousness entails that the employer, or a person acting on the employer's behalf, customises work to the abilities of the employee [18]. Quantitative components of employee obligation are specified through working time and system. According to the Labour Code, working time is not solely comprised of time of effective work. It refers to time when an employee is at the disposal of his employer to perform work duties. Therefore, what the employee sells to his employer is the ability to perform work chores in a contracted amount of time. It is up to the employer to use that time to the fullest. Being at the employer's disposal suggests a readiness to perform work duties and to follow managerial instructions. Naturally, it includes performing work duties.

The Labour Code clearly distinguishes qualitative and quantitative dimensions of working. Failure to fulfill those conditions indicates shirking. Yet, the highest normative level strongly emphasises the role of the employer, who is obliged to organise employee's work in a manner ensuring the effective use of working time, as well as achieving high efficiency and appropriate quality of work through using the employees' abilities and qualifications (art. 94 p. 2a of the L.C.). It is the employer's responsibility to plan an individual employee's work, considering his/her qualifications and abilities, so that the workload, given he/she works conscientiously and carefully, does not exceed the contracted working time. On the other hand, it is the employee's obligation to follow the supervisor's instructions as long as they are in reference to the work performed, and are not contrary to the provisions of the law and of the work contract.

4.6.2 Addressing Gap 2: Individual In-Depth Interviews to Discover Mechanisms

The main goal of the in-depth interviews was to reconstruct the process of assigning and completing work tasks in organisations. Analyses of acquired information focused on investigating semantic attributes of shirking by examining its relations to other elements in the lexicon. Therefore, six semantic networks reflecting the perceptions of what shirking is, how it comes to life and its effects were distinguished, namely: (1) equivalents, (2) oppositions, (3) employees' actions, (4) causes of shirking, (5) effects of shirking, and (6) actions towards shirking employees. Methodology for the analyses was inspired by structuralist semantics – descriptive approaches to word meaning, which assume that it cannot be studied in isolation from the relations it entertains with other expressions [19–21]. A similar approach was present in the analyses of meaning performed by [22].

Equivalents
Managers and lower-level employees see a balanced amount of not working at work as an intrinsic part of the work process, which often manifests itself in having longer breaks than the ones guaranteed by law. It provides comfort for employees, who in

turn might be more productive. Moreover, longer breaks spent on socialising and building relationships help team integration. As a result, better work outcomes may be achieved.

Oppositions

Interviewees identified two types of oppositions to a balanced amount of not working: working too much and working too little. Unsurprisingly, lower-level employees drew more attention to exploitation than managers. They often concluded that working too much is impossible and that slaves do not make good workers. It can have devastating consequences. Initially, the employees would come to work less eagerly and would be less motivated. As a result, the quality of their work would decrease. In the long run, burnout or the development of rat racing among employees could occur. Increased stress would trigger serious health consequences. Finally, employees would resign from working in harsh conditions, so staff turnover would increase.

Too little work is characterised as non-productive activities and wasted time. Managers label excessive shirkers as lazy rascals or black sheep, who intentionally procrastinate and dodge work responsibilities. Similarly, employees perceive excessive amounts of unproductive time negatively. Having nothing to do at work leaves employees wanting more chores and wishing for a better use of their time. Too little work means boredom and, in extreme cases, depression. The interviews showed that the assumption of shirking being related to higher utility than working, often made in economic theories, is false. Employees usually shirk to the extent allowed by the structure of performed work.

Employee Actions

A balanced amount of shirking is not something to openly speak about or publicly admit. It is rather an unwritten standard, which receives a certain degree of acceptance. Employees may not go the extra mile to conceal this behaviour, but they certainly feel the need to keep up appearances and not advertise it. They try not to flaunt spare time, not only because it is not looked upon favourably but also because managers assume that an employee with spare time may need more work and develop extra duties for that employee.

Causes of Shirking

Several groups of factors influencing shirking were identified by the interviewees. This section focuses only on the determinants crucial in the context of structural shirking. Most importantly, an insufficient amount of work enables shirking opportunities to arise. As managers distribute work tasks among their subordinates, a deadline for completion is established. The deadline reflects the manager's perception of how competent the employee and how difficult the task at hand. When managers undervalue employee competences they establish a long deadline, which structurally allows shirking. On the contrary, short deadlines necessitate overworking.

If the work task is well prepared and well defined, i.e. it is clear what needs to be done, employees and managers may estimate the amount of time needed to complete it with greater precision, which leaves less room for shirking. Yet, the true difficulty and complexity of work tasks is often misestimated by managers and employees. A phenomenon labelled here as the reality fluke – informational advantage of the reality over both groups of actors. Moreover, as employees increase competences, they complete tasks faster. Yet, the time for completion of a certain group of tasks cannot be decreased. This feature is labelled as task competence-dependency.

Effects of Shirking
The effects of shirking strongly depend on performed activities. Activities that result in restoration of energy, increasing knowledge or skills or strengthening connections within a network of professional contacts may increase employee productivity.

Actions Towards Shirking Employees
Managers usually turn a blind eye on balanced amounts of shirking. But when an employee crosses the line and the amount of time spent on activities other than work becomes too high, he or she is reprimanded. It is very hard to say when the employee crosses the line. Even managers were not able to define it well. They pointed to an element of reoccurrence – repetitive, longer breaks deserve a reprimand. Other forms of penalty at manager's disposal mentioned by interviewees included giving the employee more tasks to do or not promoting him/her. Drastic actions are taken very rarely. It usually takes more than just shirking to be fired. As a matter of fact, none of the interviewees has ever known or even heard of anyone being fired purely for devoting work time to unrelated activities.

The analyses allowed for answering the research questions related to the mechanisms of task performance and structural shirking. As a result, the final concept of the mechanism of structural shirking was developed.

4.6.3 *Supplementing Empirical Findings with Social Scientific Theories*

Insights from the interviews led to narrowing down the subject of interest of the second model to structural aspects of shirking in organisations, i.e. insufficient time and effort stemming from the structure of performed work. Identified assumptions from sociology of organisations, sociology of work, social psychology, management, and economics were combined into one concept of the mechanism of task performance, of which shirking is a latent consequence [23].

Interviews showed that the general structure of a work process consists of performing work tasks - role prescribed activities delegated by the manager. To perform a task, a person must (a) possess the prerequisite knowledge, (b) master the prerequisite skills, and then (c) actually choose to work on the job tasks for some period of time at some level of effort [24, p., 494]. Other factors influence task performance only indirectly, either increasing resources necessary for task performance (competence level, hereafter) or increasing motivation and subsequent level of effort. The first factor directly influencing shirking, i.e. competences, are attributes (e.g. knowledge, skills) possessed by workers, and needed for effective task performance [25]. By definition, higher competence levels increase task performance. Assumptions from social psychology also occur in economic theories. In the principal agent problem, adverse selection is regarded as private information about the quality of the agent, including information on skills, abilities, or competences [26, 27]. In the context of acquiring employment, low-skilled agents pretend to be highly skilled to receive better terms at the time of contracting. However, the reverse is the case after the contract is signed. It is in the best interest of the employee not to disclose how high his/her competences are, since it lowers his/her marginal price of labour - necessitates performing more work for the same remuneration. A large body of research showed that increasing the second direct factor influencing task performance, i.e. effort, positively affects task performance [28–31]. Effort has three main characteristics: direction, duration, and intensity [32, 33]. Such a conceptualisation of effort is coherent with the conceptualisation of working and shirking. Effort direction defines an activity as working or shirking. The two remaining attributes refer to quantitative and qualitative dimensions of working and shirking. In moral hazard, the second version of the principal-agent problem, shirking designates insufficient effort. It is in the best interest of the principal for the agent to exert a high level of effort, since it increases the probability of a high outcome. But exerting high levels of effort imposes disutility for the agent. In other words – working is costly and therefore the agent is work-averse.

In the context of information acquired during the qualitative stage of the study, research and theories from social psychology offered a coherent picture of the determinants of task performance and relations between them. Factors identified by experimental studies in social psychology, i.e. competences and effort, corresponded to assumptions present in the principal-agent problem developed in economics, i.e. adverse selection and moral hazard. The coherence was enriched by findings from sociology, where shirking is conceptualised as a latent consequence of the structure of the task performance.

4.6.4 Bringing It All Together: Implementing and Running Model #2

In the second model, shirking occurs when employees voluntarily violate organisational norms by qualitatively or quantitatively working less than expected by the manager. The two components are manifested as exerting less effort and as working less time. Shirking defines one extreme end of a continuum of aberration from manager's expectations. The opposite manifests as overworking. Both aberrations, which occur as unintended consequences of the work process implemented in the organisation, have structural components.

In a process where the manager delegates work tasks among the employees and controls their completion, employee performance (i.e. the amount of time needed by the employee to sufficiently complete a work task) is directly influenced by:

- Employee characteristics: competences and effort
- Task characteristics: difficulty and competence-dependency

In such a process, employees have two informational advantages over their managers: they know their competences (i.e. adverse selection) and exerted effort (i.e. moral hazard or qualitative shirking). As employees complete tasks delegated by their superior, they gain additional experience and increase their competence levels. Yet, reality has an informational advantage over both groups, since the actors can only guess the true task difficulty (i.e. reality fluke).

The concept of structural shirking was operationalised as a computational model that simulates an organisation comprising three employees and a set of tasks available for them to complete [34, 35]. In-depth interviews offered clear guidance for agent's decision rules. When managers delegate tasks, three aspects of their decision are always present: who? what? and when? Therefore, in the computational model a certain employee is assigned a certain task, with a clear deadline for completion. Interviews also provided a rule for task distribution – task difficulty matches employee competences and the assignment starts with the least competent employee in order to avoid a situation, in which a less competent employee is given a more difficult task than his/her more experienced colleague. As employees are completing work chores, they choose an effort level that is sufficient for a successful and timely task completion. After a certain amount of time, dependent on task difficulty, employee competence level, and employee effort, completed tasks are replaced by new chores. The interviews enabled identifying a factor of crucial importance for shirking to arise, namely the reality fluke – agents' misestimating task difficulty. Managers usually attributed it to the risk associated with the complexity of a given task – more difficult tasks are given longer deadlines and longer time reserves in case anything goes wrong. Reality, revealing the true task difficulty, surprises both the employees and their managers. As a result, the manager does not know if a failed deadline occurred due to employee shirking or due to misestimated task difficulty. Interviews also provided information on when employees notify their managers about task completion. As it turns out, it rarely takes place before the deadline. Disclosing the fact that a task was performed faster results in assigning new, more-difficult tasks earlier. Once the employee notifies the manager that the task is completed, his true and perceived competences increase. The qualitative data-gathering stage of the study also supplied assumptions for introducing learning into the model. The chosen learning function, i.e. the hyperbolic tangent – a sigmoid curve also used in neural networks, fits the descriptions of interviewees of increasing competences and decreasing amount of time needed to perform a task of a certain difficulty. The simulation ends when all employees reach proficiency in performance.

Simulations were primarily designed to test the influence of (1) initial employee competence level, (2) task difficulty misjudgement (i.e. reality fluke), and (3) initial competence misjudgement (i.e. adverse selection) on possibilities of qualitative and quantitative structural shirking. The findings showed that competence misjudgement influences only the qualitative dimension of structural shirking (i.e. the amount of exerted effort), and that the informational advantage diminishes as employees gain experience at work. However, task difficulty misjudgement, which initially influences only the amount of working time, gradually generates competence misjudgement and therefore indirectly impacts also work effort. The simulation model and results were described in detain in [34, 35].

4.7 Discussion

As described above, following the standard path of basing a computational model on theoretic assumptions did not produce the expected scientific progress. The findings of the first model were hardly surprising, thus not very informative. In the case where they were surprising, it was obvious the model did not correspond well to the real world. The shortcomings of the first model could and partially should of course be contributed to shortcomings of implementation, rather than shortcomings of the theories per se, or incompatible assumptions behind theories. However, one has to realise that theories are a reflection of their times. For instance, work process management and overview has moved from a perspective of work time regulation (e.g. the use of time stamp machines to regulate and control work quantity) to work outcome regulation.

The second model is scientifically more valuable. Not only does it offer insight into the process of task performance, it also provides clear results that are testable in experimental conditions. Developing the second model enabled a more productive inquiry into the concept of shirking, based on the analysis of legalisation surrounding labour. Labour legislation is a reflection of the processes in local organisations and of processes in the entire European Union as well as a reflection of political and ideological developments on a local, national, and supranational level. On the other hand, these dynamics have a longer time-frame than is usually of interest for a simulation model, where operationalisation of theories and correspondence to real-world processes of the short term are central, such as in our case.

More importantly, the interviews and subsequent analyses resulting in the semantic networks refocused the research problem of shirking to the process of task performance and provided insight into the mechanisms of structural shirking by identifying relevant factors and relationships between them. They also offered important guidance for agent decision rules, dynamics of agent learning process, and theoretical ranges of involved variables. Moreover, development of the second model informed us about the construction of proper indicators that should be implemented when studying the empirical reality of shirking.

Without performing qualitative analyses of legislation and inclusion of perspectives of both lower-level employees and managers any model would have had a blind spot for at least one of these players: society at large (represented and codified in legal texts), employees, and those responsible for employees' performance. This shortcoming was obvious from the development of the first model and inspired the second model, in which the mechanisms of the interplay between these levels could be studied. It is important to highlight that the meta-play between the levels, how norms are negotiated between employee and management levels, and how these are translated into and/or defined and framed by legislation remains outside of the model and is possibly an attractive subject for further research.

The use of a mixed-methods approach, enriched by carefully chosen theoretical assumptions present in social sciences, provided our story with a happy ending. The case of the model of structural shirking proves that simulation building, theory development, and the study of the empirical world are tied together in an eternal tango of research phases. An empirically grounded definition of shirking from a researcher's perspective has consequences for estimating the extent of the phenomenon, but it also has consequences for managers in preventing or making use of it as an employee in a double hermeneutic fashion.

Acknowledgements The project was supported by the National Science Centre in Poland, granted on the basis of agreement no. UMO-2014/13/N/HS6/02979.

References

1. Albanese, R., van Fleet, D.: Rational behavior in groups: the free-riding tendency. Acad. Manag. Rev. **10**(2), 244–255 (1985)
2. Demski, J., Feltham, G.: Economic incentives in budgetary control systems. Account. Rev. **53**(2), 336–359 (1978)
3. Gachter, S., Fehr, E.: Cooperation and punishment in public goods experiments. Am. Econ. Rev. **90**(4), 980–994 (2000)
4. Paulsen, R.: Empty Labor. Cambridge University Press, Cambridge (2014)
5. Holmström, B.: Moral hazard and observability. Bell J. Econ. **10**(1), 74–91 (1979)
6. Akerlof, G.A., Yellen, J.L.: Efficiency Wage Models of the Labor Market. Cambridge University Press, Cambridge (1986)
7. Shapiro, C., Stiglitz, J.: Equilibrium unemployment as a worker discipline device. Am. Econ. Rev. **74**(3), 433–444 (1984)
8. Cadsby, B., Song, F., Tapon, F.: Sorting and incentive effects of pay-for- performance: an experimental investigation. Acad. Manag. J. **50**(2), 387–405 (2007)
9. Frey, B.S.: Does monitoring increase work effort? The rivalry with trust and loyalty. Econ. Inq. **31**(4), 663–670 (1993)
10. Engel, R.P.: Why work when you can shirk?: worker productivity in an experimental setting. J. Appl. Bus. Econ. **11**(2), 104–119 (2010)
11. Judge, T.A., Chandler, T.D.: Individual-level determinants of employee shirking. Ind. Relat. **51**(3), 468–487 (1996)
12. March, J.G., Simon, H.A.: Organisations. John Wiley and Sons, New York (1958)
13. Ichino, A., Maggi, G.: Work environment and individual background: explaining regional shirking differentials in a large Italian firm. Q. J. Econ. **115**(3), 1057–1090 (2000)
14. Simon, H.A.: Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting. John Wiley and Sons, New York (1957)
15. Fudenberg, D., Tirole, J.: Game Theory. The MIT Press, Cambridge (1991)

16. Jamroży, A.: Kodeks Pracy. The Labour Code. Warszawa, C. H. Beck (2016)
17. Baran, K.W., Ćwiertniak, B.M., Driczinski, S., Góral, Z., Kosut, A., Książek, D., Wyka, T.: Kodeks pracy komentarz. Wolters Kluwer, Warszawa (2016)
18. Jaśkowski, K., Maniewska, E.: Komentarz aktualizowany do Kodeksu pracy. Wolters Kluwer, Warszawa (2016)
19. Gasparri, L., Marconi, D.: Word meaning. In: Zalta, E.N. (ed.) From The Stanford Encyclopedia of Philosophy.: http://plato.stanford.edu/archives/spr2016/entries/word- meaning/ (2016)
20. Lipka, L.: An Outline of English Lexicology. Lexical Structure, Word Semantics, and Word-Formation. Max Niemeyer Verlag, Tuebingen (1992)
21. Lyons, J.: Introduction to Theoretical Linguistics. Cambridge University Press, Cambridge (1968)
22. Robin, R.: Badanie pól semantycznych: doświadczenia Ośrodka Leksykologii Politycznej w Saint-Could. In: Głowiński, M. (ed.) Język i społeczeństwo, pp. 252–281. Czytelnik, Warszawa (1980)
23. Merton, R.K.: Social Theory and Social Structure. Free Press, New York (1968)
24. McCloy, R.A., Campbell, J.P., Cudeck, R.: A confirmatory test of a model of performance determinants. J. Appl. Psychol. **79**(4), 493–505 (1994)
25. Sandberg, J.: Competence: the basis for a smart workforce. In: Gerber, R., Lankshear, C. (eds.) Training for a Smart Workforce, pp. 47–72. Routledge, London (2000)
26. Biglasier, G., Mezzetti, C.: Principals competing for an agent in the presence of adverse selection and moral hazard. J. Econ. Theory. **61**(2), 302–330 (1993)
27. Gershkov, A., Perry, M.: Dynamic contracts with moral hazard and adverse selection. Rev. Econ. Stud. **79**(1), 268–306 (2012)
28. Blau, G.: Testing the relationship of locus of control to different performance dimensions. J. Occup. Organ. Psychol. **66**(2), 125–138 (1993)
29. Brockner, J., Grover, S., Reed, T.F., Dewitt, R.L.: Layoffs, job insecurity, and survivors' work effort: evidence of an inverted-U relationship. Acad. Manag. J. **35**(2), 413–425 (1992)
30. Gardner, D.G., Dunham, R.B., Cummings, L.L., Pierce, J.L.: Focus of attention at work: construct definition and empirical validation. J. Occup. Psychol. **62**(1), 61–77 (1989)
31. Ke, W., Zhang, P.: Motivations in open source software communities: the mediating role of effort intensity and goal commitment. Int. J. Electron. Commer. **13**(4), 39–66 (2009)
32. Bonner, S.E., Sprinkle, G.B.: The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research. Acc. Organ. Soc. **27**(4–5), 303–345 (2002)
33. Kanfer, R.: Motivation theory and industrial and organizational psychology. In: Dunnette, M.D., Hough, L.M. (eds.) Handbook of Industrial and Organizational Psychology, pp. 75–170. Consulting Psychologists Press, Palo Alto (1990)
34. Antosz, P.: The mechanism of structural shirking (Unpublished doctoral thesis). Jagiellonian University, Kraków (2017)
35. Antosz, P., Rembiasz, T., Verhagen, H.: Determinants and dynamics of individual task performance. Ergonomics (forthcoming)

Chapter 5
Causation in Agent-Based Computational Social Science

David Anzola

Abstract Even though causation is often considered a constitutive aspect of scientific explanation, agent-based computational social science, as an emergent disciplinary field, has systematically neglected the question of whether explanation using agent-based models is causal. Rather than discussing the reasons for this neglect, the article builds on the assumption that, since explanation in the field is already heavily permeated by causal reasoning and language, the articulation of a causal theory of explanation would help standardisation. With this goal in mind, the text briefly explores four candidate accounts of causation on which a causal theory of explanation in agent-based computational social science could be grounded: agent causation, algorithmic causation, interventionist causation and causal mechanisms. It suggests that, while the first two accounts are intuitively appealing, for they seem to stress the most important methodological aspects of agent-based modelling, a more robust theory of causal explanation could be developed if the field focuses, instead, on causal mechanisms and interventions.

Keywords Causation · Agent-based modelling · Social simulation · Mechanisms · Interventions

5.1 Introduction

In the philosophy of science, 'explanation' and 'causal explanation' are used interchangeably in many instances. Theories of explanation in diverse domains, including social science, consider causal explanations to be not only standard, but also the most robust form of scientific explanation (e.g. [1–4]). This conflation is understandable, given that causal reasoning seems to be connected with basic intuitions about what constitutes a good explanation, such as necessitation or

D. Anzola (✉)
Innovation Center, School of Management, Universidad del Rosario, Bogotá, Colombia
e-mail: david.anzola@urosario.edu.co

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_5

invariance [5, 6]. In agent-based computational social science (ABCSS), however, causation has been systematically neglected in the discussion about social explanation using agent-based modelling. Most references to the concept in the field are very general, e.g. through the notions of 'causal mechanism', 'causal relations' or 'causal structures', or are linked to abstract conceptual issues, e.g. downward causation.

This article does not discuss the reasons for which ABCSS has not produced a causal theory of explanation, and takes a pragmatic approach to the question of whether it should: since explanations in the field are already heavily permeated by causal reasoning and language, the articulation of a causal theory of explanation would help to standardise these explanations and make them more robust. This text takes a first step in setting the foundations for such theory by briefly exploring four candidate accounts on which a causal theory of explanation in the field could be grounded: agent causation, algorithmic causation, interventionist causation and causal mechanisms. The main claim advanced is that, while the first two accounts are intuitively appealing, for they seem to stress the most important methodological aspects of agent-based modelling, a more robust theory of causal explanation could be developed if the field focuses, instead, on causal mechanisms and interventions.

Agent and algorithmic causation have been partially discussed in ABCSS's literature; the other two accounts, causal mechanisms and interventions, have not really received much attention. The text is structured to reflect this division. The second section briefly addresses the reasons behind the call to develop a causal theory of explanation in the field. Section 5.3 discusses agent and algorithm causation, and presents limitations for both accounts. The following section, four, centres on causal mechanisms and interventions, identifying the potential advantages of developing a theory of explanation based on these accounts of causation. Some comments about the future outlook are presented last.

5.2 Why a Causal Theory of Explanation?

Practitioners of social simulation might doubt a call to develop a causal theory of explanation in ABCSS. While causal reasoning has historically permeated several explanatory accounts in social science, none of them is entirely compelling. Statistical methods, for instance, have been frequently used to represent causation mathematically. This approach, however, is often criticised because, first, it is sometimes difficult to separate causation from correlation, and, second, the identification of a causal relationship might still not lead to an illuminating explanation [7]. Functionalist accounts of explanation, in turn, become causal by reformulating the 'x is functionally relevant for y' into 'x causally relevant for y'. Several authors have also found these accounts problematic. On one hand, they are prone to swap cause and effect, making causal reasoning circular; on the other hand, they introduce an undesired teleological underpinning in the analysis of causation [8]. Finally, interpretivist accounts have incorporated causation in the analysis of the means-ends scheme. Yet, this effort has not proven successful, since some intuitions about causation such as necessitation are hard to reconcile with interpretivism. In addition, the notion of intentionality introduces some teleology into the account that is hard to test empirically [9, 10].

These three cases, naturally, do not constitute an exhaustive characterisation of social theories of causation. They, however, exemplify both the problems that have arisen in the past and further challenges posed by complexity science. Variable-centred approaches to causation, for instance, fail to capture the methodological complexity of agent-based modelling, which, as Macy and Willer [11] rightly suggest, implies a transition 'from factors to actors'. Likewise, traditional teleological explanations are not compatible with the non-centralised character of complex systems. While there is a functional component to complex phenomena, its explanatory relevance is linked to the understanding of the results of complex behaviour, e.g. self-organisation, not the causes [12]. Lastly, grounding causation on individual intention could only provide incomplete explanations of the emergence of macropatterns: complex systems are mereological structures in which there is not a direct and straightforward causal connection between levels [13]. Sometimes macropatterns end up being unintended, unexpected or conflicting when looked at from the perspective of individuals' intentions.

Beyond the question about the success of traditional accounts of social explanation, practitioners might also feel that some generalised intuitions about causation are not entirely tenable. A view of causation involving generalisation and necessitation in terms of law-like behaviour, for example, cannot be reconciled with the alleged contextual dependence of social phenomena. Likewise, the notion of linear causation, according to which effects cannot affect causes, could not accommodate the ideas of immergence or downward causation. A theory of causal explanation in ABCSS, then, would face three major obstacles: (1) traditional accounts of causal explanation in social science have not proved entirely successful, (2) complexity science poses additional challenges for causal explanation and (3) practitioners might feel some generalised intuitions about causation are misplaced or do not apply to the social domain.

These obstacles are not insignificant, yet, not insurmountable. Plus, the potential benefits from overcoming them clearly outweigh the difficulties. The most evident advantage of a causal theory of explanation is standardisation. Practitioners in the field already use a significant amount of causal reasoning and language. Concepts like 'mechanisms' or 'emergence', for example, are decidedly causal or are often depicted in causal terms. At the same time, the notion of interaction, which underlies widespread beliefs about micro to macro transition in complex social phenomena, relies on intuitions of causal production. Even the idea of generation, the most important explanatory principle in the field, is formulated using the causal terms 'sufficiency' and 'necessity'. An open discussion about causation has the potential to explicitate the criteria behind these instances of causal reasoning. Eventually, this could lead to a more robust disciplinary consensus about explanatory principles in agent-based social simulation and, subsequently, to explanatory standardisation and unification.

Addressing each obstacle in depth goes beyond the scope of this text, for that requires a more intricate philosophical discussion about causation. Before moving on, however, there are some conceptual preliminaries that are worth mentioning. Some of the most important philosophical developments in the last few decades revolve around the metaphysics of causation. Perhaps, the most significant change pertains to the popularisation of analyses of single instances of causation, under the labels of ‘actual’, ‘token’ and ‘singular’ causation [4, 14, 15]. Analyses of actual causation recognise the contextual character of causation by giving up the metaphysical search for the foundations of causation (i.e. a universal, invariable or distinctive quality of causal *relata*) to centre, instead, on how causal reasoning operates in everyday life and how scientists use models to enquire about causation.

Everyday reasoning about causation, often referred to in the literature as folk causation, is useful because it provides insights about the intuitions behind causal attribution. People introduce contextual concerns into their reasoning about causation, for example, by thinking of causation as a constrastive relationship (e.g. not as “c causes e”, but as “c causes e instead of e*”) [14] or by considering cognitive elements such as defaults (what would be the case if no additional information is provided), typicality (what is thought to be characteristic about the object or phenomenon) and normality (what is normal from a statistical or prescriptive point of view) [15]. The folk understanding of causal attribution is valuable to make sense of contextual factors, but also to dismiss relatively generalised misunderstandings (e.g. that causation is an all or nothing affair) and simplifications (e.g. that causation is a *one cause–one effect* matter). Thus, it is safe to say that, in order to develop a causal theory of explanation, ABCSS need not renounce the case-based approach to explanation that is common in model-based science, nor the presumption of context sensitivity that is typical of both social and complexity science.

The focus on what scientists do provides practitioners in the field with additional means to bypass traditional criticisms to causation. Scientists currently approach causation using different sorts of models (e.g. neuron diagrams, structural equation models, Bayes networks, etc.) that intend to reconcile the belief in an objective causal structure of the world with the context sensitivity of folk causation [16]. These models are illuminating because they elucidate causation through examples and counterexamples. By focusing on actual causation, these models are able to explore how causal intuitions could be effectively traced back to or grounded on more general principles about causation. In turn, these models differentiate between correlation and legitimate causation, for they unveil the effect of interventions, either hypothetical or real, through a causal path [4, 15]. Hence, the field can articulate a theory of causal explanation that explicitly acknowledges the influence of modelling choices in causal attributions and, at the same time, brings the interventionist nature of explanation in agent-based modelling to the forefront.

A theory of causal explanation in ABCSS needs to be developed paying attention, first, to the explanatory practices in the field and, second, to the recent developments in the philosophical literature on causation. The next two sections discuss how alternative accounts of causation in the field can, on one hand, provide more robust foundations for a theory of explanation using agent-based social simulation and, on the other hand, accommodate and make sense of recent developments in folk and scientific causation.

5.3 Existing Accounts of Causation in Agent-Based Computational Social Science

References to causation in ABCSS are scarce and usually associated with basic intuitions or platitudes about causation. The notion of causal mechanism, for example, is often used simply to acknowledge that agents in these models perform actions and that the effect of these actions can be assessed processually. Those few instances in which a more explicit connection between causation and explanation is established can be grouped in two major approaches. The first one suggests that agent-based models are causal because they enquire about the causal effects of individual action (e.g. [17]); the second links the notion of causation to the process of generation as an algorithmic feature of computational simulations (e.g. [18]).

5.3.1 Agent Causation

Using agent causation for a theory of explanation in ABCSS might, at first, seem appealing, since it would mean taking advantage of one of the most distinctive features of the method: the explicit representations of agents in the model. Agent-based modelling allows focusing on diverse physical and cognitive aspects of individual agents and their decisions, something that is uncommon or methodologically complex in variable-centred accounts of social action. In turn, it has no problems dealing with a large number of agents or spatiotemporal variations, a well-known downside of most qualitative approaches.

In spite of the methodological advantages of agent-based modelling, the focus on action might not be the most adequate alternative for a causal theory of explanation in the field. The analysis of action has significantly influenced the conceptualisation of explanation and understanding in social science. Yet, not all action-oriented accounts consider explanation should, or can, be causal [19, 20]. Those that deem causation important place the locus of causation in the connection between intention and action. This is the case both for those instances in which 'action' is approached as strategic decision-making, without emphasising much on intentionality, as well as those in which the focus is on the motivation behind the action [21, 22]. Whatever the implications this difference has for the conceptualisation of explanation, when it comes to causation, both instances strongly rely on the assumption that the social world is constituted by the consequences of individual action. This assumption is at the core of any interpretation of agent causation.

Agent causation, thus, would face a clear difficulty. While this account of causation satisfies the individualist approach to the micro-macro tension that has been adopted by most practitioners in the field, agent-based social simulation, as a method, need not provide an illuminating interpretation of causation in terms of intentions. Practitioners in the field use a very diverse set of cognitive assumptions that govern agents' decision-making. Some of these assumptions are too simple (e.g. in cellular automata) or unrealistic (e.g. zero intelligence), do not have any cognitive correlate (e.g. non-symbolic processing) or might only be meaningful when looked at from a macro perspective (e.g. biological mechanisms). These alternatives in the modelling of decision-making processes certainly permit reproducing macropatterns of interest. Yet, they do not provide any useful information about intentionality, at least in a way that is relevant to traditional action-oriented approaches of explanation.

It could be argued that agent causation could still be useful if the focus is not on providing a realistic representation of an agent's intention, but of the causal effectiveness of action. That, for example, is what Schelling seems to do when he frames the tipping model within his wider research agenda on micromotives and macrobehaviour [23]. This claim, however, ignores two major points. First, agent-based modelling could provide ontological support for effective agent causation only after a widespread agreement on the representational and material aspects of the method is achieved. Currently, the view of these aspects is far from consensual. It is not clear, for example, to what extent different agent architectures affect warrants for belief in the adequacy of a simulation. Some might believe that the field should progressively strive towards more realistic cognitive representations, whereas others might believe this is an aspect that should always be defined contextually. Likewise, while it could be argued that, in principle, there are no limitations for more accurate representations, the fact remains that computational simulation is constrained by the technological infrastructure. A practitioner might be deterred from implementing an intricate cognitive structure if, for example, there is no access to sufficient power of computation.

Thinking of explanation in terms of causal effectiveness of action is also questionable because it is conceptually reductive. The fact that agent-based models without elaborate representations of intentionality have historically provided what practitioners believe to be proper explanations raises the question about the causal and explanatory relevance of intentionality. If the output of these models is considered adequate, the structural or functional properties of interaction might trump the effects of individual action. While 'interaction' is sometimes taken as an instance of 'action oriented towards others' in social theory, that does not seem to be the case in ABCSS. Practitioners in the field frequently provide non-reductive explanation of the macropattern that emerges in the simulation. The very terms 'causal mechanism' or 'causal structure' are often used in a non-reductive manner. Furthermore, it could be argued that it is when the focus is on interaction that agent-based social simulation achieves deeper explanatory power [13]. If action is subordinated to interaction in several instances of explanation in the field, intentionality should not be given the primary causal role. This, however, goes against the basic assumption of agent causation.

Agent causation, then, is an unsuitable candidate on which to build a causal theory of explanation in the field. The focus on agents, while cashing in on a major methodological feature of the method, is problematic, for agent-based simulations are not necessarily used to enquire about the connection between intention and action. In turn, it neglects intuitions about explanation in the field, such as the relevance of interaction. Finally, it is also undermined by the lack of consensus regarding representation and the effects of the materiality of a computer simulation.

5.3.2 Algorithmic Causation

The notion of algorithmic causation in ABCSS is linked to the idea of generation. Diverse explanatory accounts in the second part of the twentieth century posit a link between causation and generation (e.g. [24–26]). The field does not inherit the concept of generation from any of these accounts, but seems to have articulated a relatively idiosyncratic approach to it. The concept of 'generation', as such, was inspired by linguistics [18]. The idea of 'generative causal explanation', however, builds upon two independent explanatory accounts. The first one links back to traditional theories of explanation, using both causal and non-causal approaches to generation (e.g. [27–30]). The other major account comes from computer science. It associates generation with computational inference or execution. Causation here is interpreted as the inferential dependence that a simulation's outcome has on the implemented model [18, 31]. Generation, then, centres on the gap between implemented model and executed simulation, and causation is mostly addressed as the possibility to inspect or backtrack macropatterns through, or after, the execution of the simulation.

Given the relevance of the methodology of computer science in the disciplinary emergence of ABCSS, the second account of generation, i.e. the one centred on computational execution, has played a more dominant role in the conceptualisation of generative causation. The account is relabelled in this text as algorithmic causation, precisely as a way to highlight its reliance on an understanding of generation that refers exclusively to its instantiation in the context of computational simulation.

Understood in this way, however, generation fails as an account of causation, because it fails, in a deeper sense, as an account of explanation. Computational simulations are epistemically opaque, i.e. the computation is too fast and intricate for the researcher to comprehend it. While inspection and verification of the execution are, in principle, possible, this is not a common practice in the field. A proper account of causation should be illuminating about the connection between cause and effect. It should also accommodate and be compatible with actual research activities. Practitioners in the field enquire about micro- to macrodynamics of emergence by focusing on the representational and experimental features of the method. Calibration and validation processes that allow for the formulation of novel knowledge claims do not have the execution of the algorithm as the main focus.

An additional difficulty with algorithmic causation is that, although the formal character of computer programming accounts for the execution of the simulation, using this algorithmic process as the causal locus could only provide understanding of causation as expectation, i.e. a simulation's step or event makes the next step expectable, according to specific values of the objects in the model and the rules of transition. Yet, expectability is a weak criterion of causation, for it fails to grasp basic aspects of widespread intuitions about causation [3, 4]. One such intuition, for example, is that causes influence or have an impact on the occurrence of the effects. In ABCSS, this notion of productivity in causal relations manifests, among other things, in the belief that the processual aspect of the simulation is causally relevant. A micro- to macrotransition is meant to be a qualitative result of the accumulated effects of interaction. Productivity, however, cannot be captured by expectability.

Even though algorithmic causation seems to highlight the processual character of a simulation, which is, arguably, one of the most distinctive methodological features of computational simulation, the concept of generation is poorly accounted for by this account of causation. The disciplinary influence of computer science has led to the emphasis of the algorithmic nature of the execution of a simulation over the epistemological issues associated with computational simulation as a dynamic type of modelling [32]. Yet, the formal character of computational simulation, by itself, is a poor foundation for a causal theory of explanation, for it is only able to yield a limited understanding of causation as expectability. While expectability could, in principle, account for some general intuitions about causation such as necessitation, the identification of causal relations through the exploration of code will probably never become a generalised practice in the field. Potential causal explanations using this account are so far removed from actual practices that they are unlikely to become illuminating.

5.4 Alternative Accounts of Causation

Both agent and algorithmic causation are poor candidates for the articulation of a causal explanatory account in ABCSS because they do not conform to current explanatory practices in the field. Causation could only be robustly introduced in the field's explanatory framework if the representational, experimental and material character of a simulation is taken into account. In what follows, it is shown how the practice of agent-based modelling can be approached as a search for causal explanations through experimental interventions on processes that can be reconstructed as mechanisms. Interventions and mechanisms are two recently developed accounts of causation with independent agendas [4, 33–35]. In some contexts, including computational social science, they have been presented as competing accounts of causal explanation [27, 36]. Yet, the two accounts could be used complementarily, for each of them aims at different aspects of the causal relationship. This complementarity becomes visible in their application to the practice of agent-based modelling.

5.4.1 Interventions

The interventionist account of causation incorporates features of manipulative and counterfactual accounts. From the former, it takes the idea that causal relationships can be unveiled by manipulation of the cause [37]; from the latter, that the semantic structure of counterfactual reasoning can capture the manipulative character of causation. Causal relationships, under this account, are not thoroughly defined metaphysically. They are simply understood as 'invariant relationships that are potentially exploitable for purposes of manipulation and control' [4, p. 17]. Interventionist accounts are called that way because they replace 'manipulation' with 'intervention'. Both notions refer to the isolation of the causal pathway from causes to effects, although 'intervention' is devised as a heuristic notion. It diverges from the traditional ideal of manipulation in that it is established counterfactually and does not rely on a reductive notion of agentic manipulation [4]. This allows for causal analysis of situations in which no intervention occurs, either human or natural, and situations in which an intervention is not possible, for example, because of moral, technical or economic reasons.

The counterfactual part of the interventionist account deals with the semantic aspect of the causal relationship, which is crucial for an account of causality in social science. One reason why causality is neglected in computational and mainstream social science is that it is usually considered to be a matter of generalisation. This is certainly a common assumption in traditional approaches to causation in social science [38, 39]. Counterfactual theories of causation, however, became popular precisely because they provide means to focus on actual causation [4, 40]. In the interventionist account, causal relationships are those that remain invariant under intervention. Invariance is a modal notion meant to identify the strength of the causal link through the identification of those circumstances in which the effect obtains, despite, or because of, the intervention [4]. In that sense, it still allows for the incorporation of contextual features.

Context has proven to be important when studying the possible extent of a generalisation in social science. While there is, for example, a large amount of literature suggesting a strong connection between economic development and democracy [41], several Latin American countries in the 1960s and 1970s followed the opposite path: dictatorships were common in countries experiencing periods of significant economic development [42]. This type of spatiotemporal specificity, far from redundant or unnecessary, is an important element to unveil the nature of this causal relationship [43]. The notion of invariance, as described above, allows incorporating the contextual features of social phenomena that are neglected by alternative accounts. It overcomes a widespread discomfort with nomological generalisation in social disciplines, while still being able to incorporate the strengths of causal reasoning.

The notion of invariance is not only useful to incorporate context, but also to demarcate both the strength and scope of the causal connection. The successful identification of the effects of an intervention is contingent on the delimitation of the counterfactual dependence and background conditions, as well as the potential sensitivity to changes in these conditions [44]. Schelling's [45] model, for example, is usually understood in terms of the effect, i.e. counterfactual relevance, of individual preferences on segregation, understood both as a dynamic and a pattern. Yet, it makes a difference for the study of preferences whether, for instance, they are analysed spatially [46], if they are taken to be a continuous or step function [47], or if they are for or against segregation [48]. If the focus is on the spatial character of segregation, counterfactual dependence will be intervened on issues such as neighbourhood size or whether relocation is made following ecological conditions. The function of the preference, which in other circumstances might be the locus of intervention, could be taken in this case as a background condition.

Interventionist theories of causation are not really mentioned in the agent-based social simulation literature. Practitioners, however, usually adopt this type of reasoning when providing explanations. The concept of 'what-if' questions is used when referring to the *possible worlds* that could be accessed with the parametric exploration of the simulation. The exploration is accounted for by the manipulative component of the intervention; the 'what if', by the counterfactual. The difference is in the framing of the explanation. 'What-if' questions are formulated centring on the initialising conditions, leaving the output unaddressed. Counterfactuals, conversely, explicitly link initialising conditions and results. In the specific case of agent-based modelling, the cognitive difference between 'what-if' questions and counterfactuals is that they are formulated before and after the execution of the simulation, respectively.

Practitioners of agent-based social simulation can more explicitly use interventionist theories of causation to frame their explanation within a causal account that is able to incorporate the contextual nature of social phenomena and the experimental underpinning of the method. They can also do it in a way that bypasses the most common criticisms of this approach to causation. Agent-based modelling provides three main advantages when exploring the computational model through causal interventions. The first one is linked to *the who that performs the intervention*. In comparison with real experiments, using artificial societies makes all interventions, in principle, possible. In turn, the negatively valued anthropocentric character of traditional accounts of manipulative causation is also absent, since, for the artificial system, the researcher is in the position of an omnipotent observer. A second advantage is linked to *the who that is affected by the intervention*. Manipulative accounts have been questioned because the agents intervened can be reflexive about the intervention, thus affecting the outcome [49, 50]. Yet, artificial agents lack the kind of reflexivity that could hamper the identification of causal effects. A final advantage pertains to *the context of intervention*. Manipulative accounts have been criticised because there is never certainty about whether the causal pathway between the cause and the effect has been entirely isolated [49, 51, 52]. The modular character of computer programming, however, provides a way to guarantee this isolation in the structural and functional features of a simulation, making it amenable for controlled manipulation. Agent-based modelling provides a context in which interventions could be explored without the difficulties that real-world situations pose for control.

For a theory of causal explanation based on interventions to work properly, however, the field needs to detach itself from the algorithmic focus inherited from computer science. There are some current practices in ABCSS that conflict with the representational and experimental understanding of agent-based modelling, such as approaching validation as benchmarking or the lack of clarity regarding the epistemological status of calibration. By reorganising practices to bring modelling and experimentation to the forefront, practitioners can better use agent-based models to unveil causal relations through the identifications of counterfactual implications of structural and parametric modifications. At the same time, an explicit account of causation could shed light on how representation operates in the field and how folk and scientific intuitions about causation manifest in an indirect approach to knowledge, such as modelling.

5.4.2 Mechanisms

An additional advantage of agent-based modelling is that it provides a way to explain counterfactual dependence processually. Intuitions about the explanatory and causal relevance of the processual character of a simulation can be better accounted for by linking the agenda of ABCSS with the agenda of contemporary mechanism in general philosophy of science. Contemporary mechanism is a processual account of causal explanation, based on the analysis of 'mechanism' as '[...] entities and activities organised such that they are productive of regular changes from start or set-up to finish or termination conditions' [34, p. 3].

There are two fundamental elements in the mechanist account that are worth mentioning: the notions of 'activity' and 'production'. The first one is meant to set mechanisms apart from static entity-based approaches to explanation, in which processes are reduced to changes in properties. This is done by implementing a double ontology of entities and activities [53–55]. The former concept is understood in a relatively straightforward manner and the latter as type of causes or producers of change. In social science, activities could be associated with nominalised verbs, such as socialisation or structuration. The key to this ontological category is that it transcends reductive causal explanation, i.e. activities cannot be reduced to individuals and their actions. The more intricate metaphysical base of agent-based models, which includes human and non-human entities, combined with the dynamic aspect of the simulation, can be used to explore this double ontology and, in particular, the role of activities.

The concept of activities is put forward by contemporary mechanism to bring to the forefront the notion of production. This notion is particularly relevant in folk accounts of causation, for it is something experienced by regular people in everyday life. It has to do with the possibility to capturing the manifestation of change in the phenomenon of interest, usually through perception and language: it is evidenced, for example, by verbs that describe causal relations, e.g. 'I *moved* the table'. Production has not been properly articulated into traditional causal accounts, on one hand, because of the difficulties to identify a generalised loci of production that does not depend on perception and language and, on the other hand, because processes are usually accounted for by probability-based state transitions [5, 56, 57]. Yet, since the contemporary analysis of causation renounces causal primitivism and focuses instead on what scientists do when they posit causal relations, production and its connection with causation have received a renewed attention [4, 16, 57]. In turn, since agent-based modelling has a representation of the process, i.e. the simulation is meant to account for the temporal evolution of the system, the method can provide an explicit account of production. Overall changes during the simulation can be linked to actual interaction instead of probability. Because agent-based models are temporal models, there is an explicit exploration of the causal path.

From the point of view of explanation, causal production is the key element to understand the role of emergence in complex social dynamics. It provides the basis for a non-reductive approach both to the macro level and to social interaction dynamics. It also brings to the fore complexity science's concerns with trajectories, reflected, for example, in concepts such as sensitivity to initial conditions, path-dependence and non-linearity. The embeddedness of interaction in agent-based modelling, in combination with the experimental character of computational simulation, can make the method an excellent tool for the analysis of causation. By focusing on interventions and mechanisms, agent-based modelling could help to bridge the gap between the two major approaches to causation: difference-making and causal processes [58–61]. To do this, however, some methodological changes are required. For example, causal mechanical explanation works at its best when the process is fully acknowledged. In agent-based social simulation, however, attention is most often paid to the initial and final conditions, mostly because of how traditional verification and validation techniques operate, especially those transferred from the quantitative domain.

There is a complex relationship between explanation, causation and validation in ABCSS, for knowledge claims in the field rely on the use of models as surrogates for thinking. Grüne-Yanoff [62], for example, suggests computer simulations cannot provide causal explanations in social science because they cannot be completely validated. Yet, it is not the validation of a simulation that determines its causal explanatory relevance, but its contribution to, and accommodation of, the knowledge about the counterfactual dependence observed in a particular phenomenon. The research programme of segregation dynamics constitutes a paradigmatic example. There is a robust research programme on the analysis of segregation as an abstract dynamic in social and spatial dimensions [47, 63]. Agent-based models help to render explanations of this phenomenon illuminating by providing, first, a formal experimental setting in which the basic features of self-reinforcing segregation processes can be tested and, second, processual evidence of the clustering dynamics, through direct inspection of the model's visualisation or diverse data-generation processes. Practitioners do not use causal reasoning simply for the validation of the models, but also for their design and operation.

While interventions allow incorporating the experimental nature of agent-based modelling, mechanisms offer a mean to satisfy those intuitions that practitioners have about the causal relevance of generation when using agent-based social simulation. These intuitions, which were poorly accounted for by algorithmic causation, are at the core of the field's foundational narrative regarding the emergence, self-organisation and adaptiveness of complex social phenomena. Intuitions about generation are, in part, the reason for which the field is thought to provide non-reductive explanations that could be framed within the wider agenda of complexity science.

5.5 Conclusion

This text discussed the merits and disadvantages of four candidate accounts of causal explanation in ABCSS. The first two accounts, agent and algorithmic causation, were found lacking, for they do not conform to the practices and research agenda of the field. Agent causation requires an emphasis on intentionality that the field does not have. In turn, algorithmic causation could only yield a weak understanding of causation as expectation, while unnecessarily downplaying the representational and experimental features of agent-based modelling. Conversely, the other two accounts, causal mechanisms and interventions, were considered to, first, satisfy general intuitions about explanation in the field, second, adequately accommodate recent developments in the philosophical literature on causation and, third, bring to the forefront the experimental and representational features of the method.

For a sound causal theory of explanation to be developed in the field, a few conceptual and methodological challenges should be overcome. Some challenges that are directly related to the understanding of interventions and mechanism were briefly mentioned; there are others, however, that were not addressed because they are intertwined in a more complex way with the overall practice of agent-based social simulation. First, it is clear that there is a connection between the intricateness of the model and the possibility to unveil causal connections. Intricate models, even with the modular advantage provided by computer simulation, can make it impossible to identify these connections. Second, practitioners have to devise tools and techniques to understand and measure the dynamics, so as to reduce the epistemic opacity of computational simulation. Counterfactual understanding of the process can be increased by the implementation of simple measures, such as longitudinal indexes and indicators. Other options, such as the visual exploration of the simulation's response to interventions while the model is running, can be of help. Finally, there needs to be an inquiry about how representation directly affects warrants for belief during the process of validation. Given that knowledge is produced indirectly, there is a widespread consensus about the importance of empirical data and theory. In spite of this consensus, the field has yet to understand how knowledge claims associated with external data and theory are permeated and affected by different approaches to modelling and representation.

When dealing with these challenges, practitioners need not invest time trying to successfully answer all of the traditional criticisms to causal explanation. There is no pressing need to identify the philosophical features of the causal structure underlying complex social phenomena or to look for causal primitives upon which to build a generalised account of causation. ABCSS should aim towards developing a causal theory of explanation that, first, focuses on how practitioners of agent-based social simulation articulate causal models of explanations and, second, is non-reductive. By doing so, the field could strengthen its account of explanation, while avoiding some of the pitfalls that have wrongly led to question the connection between causation and explanation.

References

1. Elster, J.: Explaining Social Behavior. Cambridge University Press, New York (2007)
2. Hedström, P.: Dissecting the Social. Cambridge University Press, New York (2005)
3. Salmon, W.: Causality and Explanation. Oxford University Press, Oxford (1998)
4. Woodward, J.: Making Things Happen. Oxford University Press, New York (2003)
5. Danks, D.: The psychology of causal perception and reasoning. In: Beebee, H., Hitchcock, C., Menzies, P. (eds.) The Oxford Handbook of Causation. Oxford University Press, Oxford (2010)
6. Woodward, J.: Causal perception and causal cognition. In: Roessler, J., Lerman, H., Eilan, N. (eds.) Perception, Causation, and Objectivity. Oxford University Press, Oxford (2011)
7. Freedman, D.: From association to causation: some remarks on the history of statistics. Stat. Sci. **14**, 243–258 (1999)
8. Turner, S.: Cause, the persistence of teleology, and the origins of the philosophy of social science. In: Turner, S., Roth, P. (eds.) The Blackwell Guide to the Philosophy of the Social Sciences. Blackwell, Oxford (2003)
9. Sehon, S.: Goal-directed action and teleological explanation. In: Campbell, J., O'Rourke, M., Silverstein, H. (eds.) Causation and Explanation. MIT Press, Cambridge, MA (2007)
10. Sintonen, M.: Explanation: in search of the rationale. In: Kitcher, P., Salmon, W. (eds.) Scientific Explanation. University of Minnesota Press, Minneapolis (1989)
11. Macy, M., Willer, R.: From factors to actors: computational sociology and agent-based modeling. Annu. Rev. Sociol. **28**, 143–166 (2002)
12. Anzola, D., Barbrook-Johnson, P., Cano, J.: Self-organization and social science. Comput. Math. Organ. Theory. **23**, 221–257 (2017)
13. Anzola, D.: The philosophy of computational social science. Ph.D. thesis, Department of Sociology, University of Surrey (2015). http://epubs.surrey.ac.uk/id/eprint/808102
14. Schaffer, J.: Causal contextualism. In: Blaauw, M. (ed.) Contrastivism in Philosophy. Routledge, New York (2012)
15. Halpern, J., Hitchcock, C.: Graded causation and defaults. Br. J. Philos. Sci. **66**, 413–457 (2015)
16. Hitchcock, C.: Three concepts of causation. Philos. Compass. **2**, 508–516 (2007)

17. Tesfatsion, L.: Agent-based computational economics: a constructive approach to economic theory. In: Tesfatsion, L., Judd, K. (eds.) Handbook of Computational Economics. Elsevier, London (2006)
18. Epstein, J.: Agent-based computational models and generative social science. Complexity. **4**, 41–60 (1999)
19. Joas, H.: The Creativity of Action. University of Chicago Press, Chicago (1996)
20. Stones, R.: Theories of social action. In: Turner, B. (ed.) The New Blackwell Companion to Social Theory. Wiley-Blackwell, New York (2009)
21. Helle, H.: The classical foundations of micro-sociological paradigm. In: Helle, H., Eisenstadt, S. (eds.) Micro Sociological Theory. Sage, Bristol (1985)
22. Winch, P.: The Idea of Social Science and Its Relation to Philosophy. Routledge, London (1990)
23. Schelling, T.: Micromotives and Macrovehavior. W.W. Norton & Co., New York (1978)
24. Bhaskar, R.: The Possibility of Naturalism. Routledge, London (1998)
25. Bunge, M.: How does it work?: the search for explanatory mechanisms. Philos. Soc. Sci. **34**, 182–210 (2004)
26. Harré, R.: The Philosophies of Science. Oxford University Press, Oxford (1985)
27. Elsenbroich, C.: Explanation in agent-based modelling: functions, causality or mechanisms? J. Artif. Soc. Soc. Simul. **15**, (2012)
28. Hedström, P., Ylikoski, P.: Causal mechanisms in the social sciences. Annu. Rev. Sociol. **36**, 49–67 (2010)
29. Manzo, G.: Variables, mechanisms, and simulations: can the three methods be synthesized?: a critical analysis of the literature. Rev. Fr. Sociol. **48**, 35–71 (2007)
30. Squazzoni, F.: Agent-Based Computational Sociology. Wiley, London (2012)
31. Doreian, P.: Causality in social network analysis. Sociol. Methods Res. **30**, 81–114 (2001)
32. Anzola, D.: Knowledge transfer in agent-based computational social science. Stud. Hist. Philos. Sci. Part A. **37**, 29–38 (2018). https://doi.org/10.1016/j.shpsa.2018.05.001
33. Bechtel, W., Abrahamsen, A.: Explanation: a mechanist alternative. Stud. Hist. Philos. Sci. Part C Stud. Hist. Philos. Biol. Biomed. Sci. **36**, 421–441 (2005)
34. Machamer, P., Darden, L., Craver, C.: Thinking about mechanisms. Philos. Sci. **67**, 1–25 (2000)
35. Woodward, J.: Counterfactuals and causal explanation. Int. Stud. Philos. Sci. **18**, 41–72 (2004)
36. Waskan, J.: Mechanistic explanation at the limit. Synthese. **183**, 389–408 (2011)
37. Menzies, P., Price, H.: Causation as a secondary quality. Br. J. Philos. Sci. **44**, 187–203 (1993)
38. Barringer, S., Eliason, S., Leahey, E.: A history of causal analysis in the social sciences. In: Morgan, S. (ed.) Handbook of Causal Analysis for Social Research. Springer, Berlin (2013)
39. Goertz, G., Mahoney, J.: A Tale of Two Cultures. Princeton University Press, Princeton (2012)
40. Lewis, D.: Counterfactuals. Blackwell, London (2001)
41. Robinson, J.: Economic development and democracy. Annu. Rev. Polit. Sci. **9**, 503–527 (2006)
42. O'Donnell, G.: Modernization and Bureaucratic-Authoritarianism. University of California, Berkeley, Institute of International Studies (1979)
43. Mainwaring, S., Pérez-Liñán, A.: Level of development and democracy: Latin American exceptionalism, 1945-1996. Comp. Polit. Stud. **36**, 1031–1067 (2003)
44. Woodward, J.: Sensitive and insensitive causation. Philos. Rev. **115**, 1–50 (2006)
45. Schelling, T.: Dynamic models of segregation. J. Math. Sociol. **1**, 143–186 (1971)
46. Reardon, S., O'Sullivan, D.: Measures of spatial segregation. Sociol. Methodol. **34**, 121–162 (2004)
47. Bruch, E., Mare, R.: Segregation dynamics. In: Hedström, P., Bearman, P. (eds.) The Oxford Handbook of Analytical Sociology. Oxford University Press, Oxford (2009)
48. Macy, M., Centola, D., Flache, A., van de Rijt, A., Willer, R.: Social mechanisms and generative explanations: computational models with double agents. In: Demeulenaere, P. (ed.) Analytical Sociology and Social Mechanisms. Cambridge University Press, Cambridge (2011)
49. Goldthorpe, J.: Causation, statistics, and sociology. Eur. Sociol. Rev. **17**, 1–20 (2000)

50. Woodward, J.: Interventionist theories of causation in psychological perspective. In: Gopnik, A., Schulz, L. (eds.) Causal Learning. Psychology, Philosophy, and Computation. Oxford University Press, New York (2007)
51. Bogen, J.: Analysing causality: the opposite of counterfactual is factual. Int. Stud. Philos. Sci. **18**, 3–26 (2004)
52. Woodward, J.: Causal models in the social sciences. In: Turner, S., Risjord, M. (eds.) Philosophy of Anthropology and Sociology. Elsevier, Amsterdam (2007)
53. Glennan, S.: Mechanisms. In: Beebee, H., Hitchcock, C., Menzies, P. (eds.) The Oxford Handbook of Causation. Oxford University Press, Oxford (2010)
54. Illari, P., Williamson, J.: In defence of activities. J. Gen. Philos. Sci. **44**, 69–83 (2013)
55. Tabery, J.: Synthesizing activities and interactions in the concept of a mechanism. Philos. Sci. **71**, 1–15 (2004)
56. Anscombe, G.: Causality and determination. In: Sosa, E., Tooley, M. (eds.) Causation. Oxford University Press, Oxford (1993)
57. Illari, P.: Why theories of causality need production: an information transmission account. Philos. Technol. **24**, 95–114 (2010)
58. Campaner, R., Galavotti, M.: Some remarks on causality and invariance. In: Arturo, C. (ed.) Causality, Meaningful Complexity and Embodied Cognition. Springer, Berlin (2010)
59. Craver, C.: Explaining the Brain. Oxford University Press, Oxford (2009)
60. Woodward, J.: Mechanisms revisited. Synthese. **183**, 409–427 (2011)
61. Woodward, J.: Mechanistic explanation: its scope and limits. Aristot. Soc. Suppl. Vol. **87**, 39–65 (2013)
62. Grüne-Yanoff, T.: The explanatory potential of artificial societies. Synthese. **169**, 539–555 (2008)
63. Fossett, M., Dietrich, D.: Effects of city size, shape, and form, and neighborhood size and shape in agent-based models of residential segregation: are schelling-style preference effects robust? Environ. Plan. B Plan. Des. **36**, 149–169 (2009)

Chapter 6
Times of Crises and Labour Market Reforms

An ABM Evaluation

Tom Bauermann

Abstract This work assesses the effectiveness of labour market policies in reducing unemployment during a recession. In order to do so, I build an agent-based model and analyse the impact of typical labour market policies on the short-, medium- and long-term development of unemployment. I analyse the effect of (1) a reduction of unemployment benefits (UB), (2) increasing search efforts among unemployed, (3) a mixture of both aforementioned policy responses, (4) governmental transfers to steer consumption and (5) short-time work. In the following extended abstract, I present the results of two policy experiments, (3) and (5). The methods I use here are slightly different from those in the existing ABM literature. The analyses are carried out as policy experiments. In order to evaluate the effectiveness of a labour market policy, the graphics of this work illustrate range of effects of the specific policy measure on unemployment (mean and confidence intervals across several runs) instead of solely depicting the mean development. In contrast to equilibrium models, I find that measures which raise the pressure on unemployed to find a job, measures (1) to (3), have no noteworthy impact or slightly increase unemployment in the long run. Instead, governmental interventions which stabilise aggregate supply and demand, like (4) and (5), rather dampen unemployment.

Keywords Agent-based model · Labour market policies · Macroeconomics · Unemployment · Behavioural economics

T. Bauermann (✉)
Chair of Macroeconomics, Ruhr-University Bochum, Bochum, Germany

Ruhr-Graduate School in Economics, RWI – Leibniz Institute for Economic Research, Essen, Germany
e-mail: tom.bauermann@ruhr-uni-bochum.de

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_6

6.1 Introduction and (Short) Model Description

The outbreak of the Great Recession provoked governments around the world to respond to a growing number of unemployed in the respective countries. For instance, a standard policy measure was altering the generosity of unemployment benefits (UB). Due to their use, labour market measures became a topic of intensive debate in macroeconomics again. Studies, using equilibrium models to analyse measures like the extension of UB, find mainly adverse effects on unemployment, e.g. [11]. However, equilibrium models disregard the aggregate demand and model the recession as a drop in productivity. Thus, the effects of policy interventions in response to other reasons for a recession, e.g. firm insolvencies, and on aggregate demand have not been analysed. By applying an agent-based modelling approach (ABM), this model can overcome the mentioned inadequacies of mainstream models. For example, due to the implied heterogeneity of firms, recessions are usually caused by firm insolvencies instead of productivity shocks. This further implies that, instead of hardly interpretable exogenous shocks, fluctuations in ABM emerge endogenously from the decisions of agents. Therefore, ABM responds to Romer's critique [13]. Another advantage, demand for goods in an ABM, is usually, endogenously determined by its agents. Thus, ABM can show how labour market policies affect unemployment simultaneously via different channels, e.g. via the impact on consumption. The (full) paper analyses the effects and mechanisms of (1) a reduction of unemployment benefits (UB), (2) increasing search efforts among unemployed, (3) a mixture of both aforementioned policy responses, (4) governmental transfers to steer consumption and (5) short-time work. Here, I only analyze measure (3) and (5).

The model which is used in this work is a modification of the baseline model of Delli Gatti and co-authors [2] (henceforth: DG2011). In this extended abstract, I only describe the major modifications.[1] While the model of DG2011 contains firms, workers and banks, this work introduces a government and a central bank. The model is stock-flow consistent by following the logic of [6]. The government collects taxes and pays transfers to the banking sector, firms and unemployed. Also modified, workers determine consumption and saving by applying the concept of buffer-stock saving, only unemployed search for employment and they have a reservation wage which is determined by markups over UB [14]. In comparison to the baseline model of DG2011, another modification is related to exits and entries of firms. As it is standard in Macro-ABM, e.g. in [2], every market exit is replaced by an entry such that the number of firms is constant. In the mentioned DG2011, new firm's wealth and prices are copied from a small incumbent firm and a number of unemployed is automatically added to the entrant. In this model, the government provides funds in form of start-up subsidies, newcomers apply a low price strategy

[1]For the model code, please contact the author: tom.bauermann@ruhr-uni-bochum.de.
An early version of the full paper can be downloaded from: https://www.boeckler.de/pdf/v_2018_10_26_bauermann.pdf.

and a search process of the new firms replaces the automatic matching of workers. I assume that new firms avoid losses by also offering relatively low wages. The setting w.r.t. entries and exits, which was made up for this work, follows empirical and theoretical investigations. Entries and exits are highly correlated in most of the Western economies [5]. Another advantage, the one-on-one assumption, avoids monopolisation (see: [2]). Based on [12], new firms which enter the market without entry barriers do usually not compete on technology, and these newcomers offer low-wage contracts.

6.2 Results

The policy measures (1) to (5) are analysed similar to impulse response functions in the mainstream literature in order to take into account the uncertainty regarding their effects. The policy responses are analysed as follows: When the economy is in the middle of a recession, indicated by a predefined unemployment level that exceeds the average unemployment level,[2] the government is triggered to take up a measure to fight unemployment (policy shift). After the policy measure occurred, the model is simulated for the following 40 periods to illustrate the development of unemployment over time. Each period marks one quarter in real time. I simulate the time series with (policy scenario) and without the policy measure (baseline scenario), each for 700 runs. The described procedure allows me to compute and illustrate the mean development of unemployment with and without a policy measure. The latter is, so to say, the countefactual. In addition, I compute and depict the mean development of the (relative) difference between both mentioned unemployment paths (policy and baseline scenario) for the above-mentioned following 40 periods. Thus, the effect of the policy, compared to the absence of a governmental intervention on the labour market, becomes visible. Finally, I compute the standard errors to construct the confidence bands surrounding the mean difference between policy and baseline scenario. Since ABM relies on stochastic processes, the illustration of confidence intervals is necessary to examine whether the (potential) gaps in the development of unemployment between the baseline and the policy scenario are systematic or due to randomness. Thus, due to the explained procedure, the (in-)effectiveness of a policy instrument becomes visible if the mean difference is (in-)significantly different from zero.[3]

---

[2]The average unemployment in the model is at 9.01% and its average standard deviation 2.0%. The government starts its intervention when unemployment exceeds 13.0% or, in other words, when unemployment exceeds mean unemployment by 2*SD.

[3]To determine the number of runs (i.e. 700 random seeds), I follow the concept of variance stability [10]. The variance of the variates settles to the mean variance, 700 runs. To ensure that this burn-in phase was passed when the government starts its policy, the policy shift starts at period 1000. The 95% confidence bands follow the standards in the literature [3].

As described above, the simulation exercise is performed for five different policy experiments (policy shifts). This extended abstract presents only two of them, a simultaneous increase in search efforts of unemployed and reduction in UB and short-time work. It should be noticed that the model has a long-run (statistical) equilibrium unemployment rate [7]. Therefore, even without a policy shift, unemployment returns to its equilibrium rate after a certain period of time. Hence, I analyse which measure accelerates the drop in unemployment and whether it returns to the normal unemployment rate, given the policy intervention. The first policy experiment analyses the impact of a permanent increase in the number of applications sent per unemployed from three to six and a permanent decrease of UB from 50% to 30% of the average wage.[4] The measures may increase the number of hirings and, thus, decrease unemployment due to two reasons. In terms of this model, first, a cut in UB leads to a reduction of the reservation wages of unemployed which (may) makes it easier for new firms to find workers by offering low wages. Second, increasing the number of applications may lead to a reduction of coordination failures, the coexistence of open vacancies and unemployed willing to fill these vacancies. However, following Fig. 6.1, the attempt to reduce unemployment by changing the behaviour of unemployed does not lead to an accelerated decrease in unemployment. The left plot illustrates that unemployment does not differ between the baseline and the policy scenario, at least in the first 20 periods. The right graph confirms this view, and, by looking at the confidence bands, the difference between the scenarios is rather insignificant. Following the left graph, unemployment even exceeds the normal rate as time evolves. These developments can be explained as follows: Both measures, indeed, decrease the duration of unemployment, at least in the short run. However, since coordination failures are negligible during recessions, because posted vacancies are rather low, and since the reduction of UB has negative effects on consumption and, thus, on labour demand, the positive effects on the labour market are very small and outweighed by the negative effects. Hence, unemployment rather increases over time instead of returning to its equilibrium rate. Nevertheless, it should be noted that the relative difference deviates insignificantly from zero. Performing the experiment for both measures individually also leads to insignificant differences. These results are confirmed by other studies. For instance, in an empirical work, [9] can neither detect a clear correlation nor a clear causation between the welfare system and unemployment.

In the second policy experiment, the government gives firms the opportunity to implement short-time work. The programme is kept for four periods. Firms

[4]The baseline ratio of unemployment benefits to average wage, 50%, is, roughly, the average net replacement rate (NRR) in Germany in the early 2000s, which was taken from the NRR statistics [15]. The magnitude of the reduction was found by following the reduction of UB of long-term unemployed in Germany in the course of the labour market reforms in the 2000s. The reduction of this work is a bit milder compared to the mentioned reduction [4]. Due to the absence of a calibration criteria for search efforts, doubling the search efforts mimics a significant but still realistic increase.

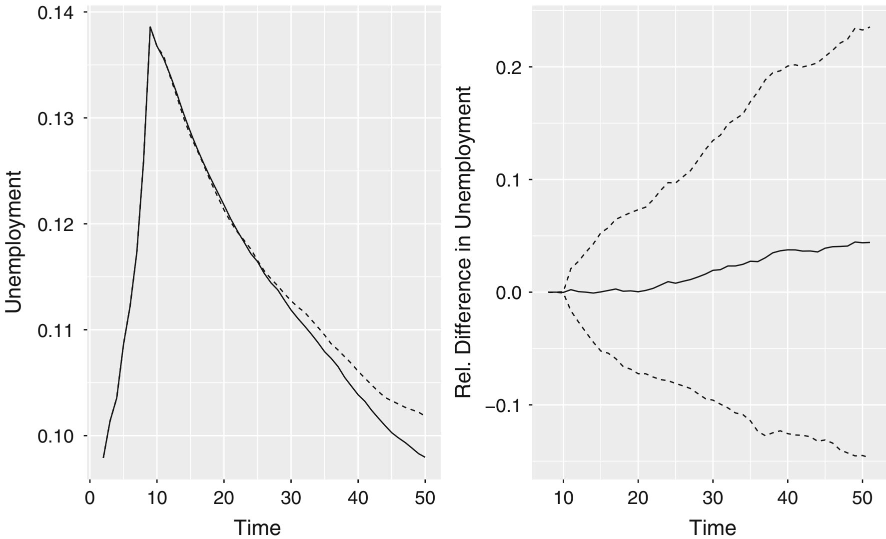


Fig. 6.1 Plots display the effect of an increase in search efforts and a reduction of UB (simultaneous policy shift). The left plot displays the mean development of unemployment for baseline (solid) and policy scenario (dashed). The shift starts in period 8. The right plot displays the average relative difference between baseline and policy scenario (solid) and its 95%-confidence interval (dashed). Here, the shift starts in period 3

with positive but shrinking wealth can reduce output by reducing hours worked proportionally and, simultaneously, the wage bill. The government partly compensates workers for the loss in income by providing 60% of the net income loss.[5] As shown in Fig. 6.2, short-time work reduces unemployment quickly (left graph) and systematically (right graph). Further, Fig. 6.2 shows that the unemployment rate converges to the equilibrium level in the longer run. Thus, the taken policy response accelerates the reduction of unemployment but does not change the normal rate permanently. The effect can be explained as follows: On the one hand, since short-time work implies a reduction of output, firms become more reluctant to lay off workers to reduce losses. In consequence, also a massive drop in consumption is avoided (demand effect). On the other hand, firms that do not apply short-time work face a spill-over effect since they can use the “free space” on the goods market given by the firms which apply short-time work (supply effect). Both types of firms stabilise their net worth, which improves their expectations and therefore labour demand. Such a relaxation on the goods market helps, particularly, if competition is fierce due to a lack of demand, e.g. a recession. Besides short-time work, a unique, debt-financed extraordinary transfer to workers, policy experiment (4),

[5]The short-time work scheme followed, roughly, the German short-time work in the aftermath of the recent financial crisis [1]. The firms could reduce hours worked and wage bill down to 45%.

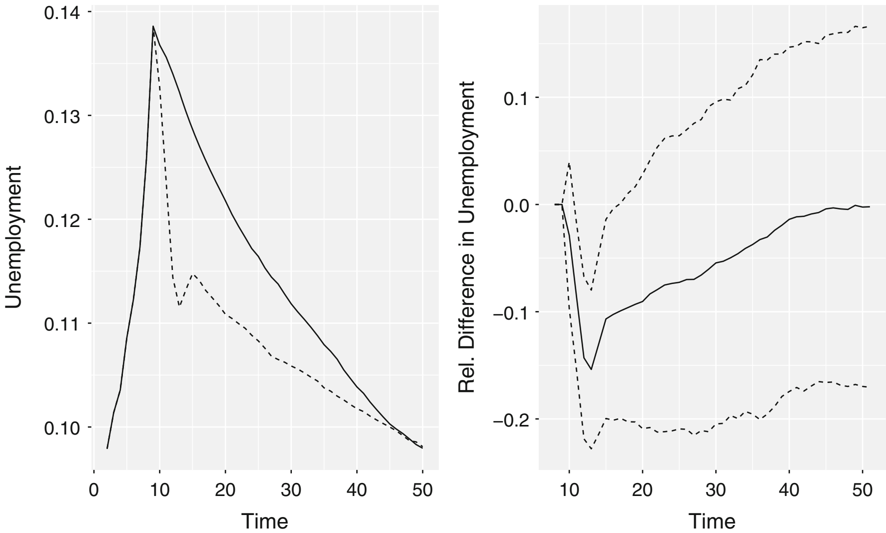


Fig. 6.2 Plots display the effect of an introduction of short-time work. The left plot displays the mean development of unemployment for baseline (solid) and policy scenario (dashed). The shift starts in period 8. The right plot displays the average relative difference between baseline and policy scenario (solid) and its 95% confidence interval (dashed). Here, the shift starts in period 3

yields similar results. The described effects of short-time work find support by other works. Empirical studies find strong support that short-time work had a dampening effect on unemployment during the crisis (e.g. [8]).

References

1. Brenke, K., Rinne, U., Zimmermann, K.F.: Short-time work: the German answer to the great recession. Int. Labour Rev. **152**, 287–305 (2013)
2. Delli Gatti, D., Desiderio, S., Gaffeo, E., et al.: Macroeconomics from the Bottom-Up. Springer, Milano (2011)
3. Dosi, G., Napoletano, M., Roventini, A., et al.: Rational heuristics? Expectations and behaviors in evolving economies with heterogeneous interacting agents. LEM Work. Paper Ser. **31**, 1–41 (2017)
4. Engbom, N., Detragiache, E., Raei, F.: The German labor market reforms and post-unemployment earnings. IMF Work. Paper **15**(162), 1–26 (2015)
5. Geroski, P.: What do we know about entry? Int. J. Ind. Organ. **13**, 421–440 (1995)
6. Godley, W., Lavoie, M.: Monetary Economics. An Integrated Approach to Credit, Money, Income, Production and Wealth, 2nd edn. Palgrave Macmillan, Hampshire (2012)
7. Grazzini, J., Richiardi, M.: The agent-based experiment. In: Delli Gatti, D., Fagiolo, G., Gallegati, M., et al. (eds.) Agent-Based Models in Economics: A Toolkit. Cambridge University Press, Cambridge (2018)

8. Herzog-Stein, A., Lindner, F., Sturn, S.: The German employment miracle in the Great Recession: the significance and institutional foundations of temporary working-time reductions. Oxf. Econ. Pap. **70**, 206–224 (2018)
9. Howell, D., Rehm, M.: Unemployment compensation and high European unemployment: a reassessment with new benefit indicators. Oxf. Rev. Econ. Policy **25**, 60–93 (2009)
10. Lee, J.S., Filatova, T., Ligmann-Zielinska, A., et al.: The complexities of agent-based modeling output analysis. J. Artif. Soc. Soc. Simul. **18**, 1–26 (2015)
11. Mitman, K., Rabinovich, S.: Unemployment benefit extensions caused jobless recoveries!? PIER Working Paper 14-013:1-33 (2014)
12. Nystroem, K., Elvung, G.Z.: New firms and labor market entrants: is there a wage penalty for employment in new firms? Small Bus. Econ. **43**, 399–410 (2014)
13. Romer, P.: The Trouble with Macroeconomics. Available via Personal Homepage. https://paulromer.net/wp-content/uploads/2016/09/WP-Trouble.pdf (2016). Cited 02 May 2018
14. Salle, I., Seppecher, P.: Deleveraging crises and deep recessions: a behavioural approach. Appl. Econ. **47**, 3771–3790 (2015)
15. van Vliet, O., Caminada, K.: Unemployment replacement rates dataset. https://www.universiteitleiden.nl/en/law/institute-for-tax-law-and-economics/economics/data-sets/unemployment-replacement-rates-dataset (2018). Cited 02 May 2018

Chapter 7
Selecting the Right Game Concept for Social Simulation of Real-World Systems

Femke Bekius and Sebastiaan Meijer

Abstract Game theoretical models can be used for social simulation of real-world systems. These models describe the interaction between actors who have to make a decision. Before application of those models, the game concept that describes the situation at hand needs to be selected. Selecting the *right* game concept is crucial when choosing a model to simulate the process, but not trivial. In this paper we present a taxonomy of game concepts to select a set of suitable models.

Keywords Game theory · Decision-making · Game taxonomy · Actor behavior · Social simulation

7.1 Introduction

Decision-making on real-world complex adaptive systems (CAS) is complex due to the technical uncertainties of the system. Examples of CAS are railway systems and healthcare systems. Both consist of several interdependent systems which interact and need to work together to let the entire system function [22, 30]. The sum of the interdependent systems in a CAS is more than just the sum of its parts. CAS show emerging behavior, i.e., they evolve and adapt over time based on new information [23].

Simulating the system using different scenarios creates insight in the uncertainties and behavior of the system [29]. Social simulation of real-world systems mainly focuses on sociotechnical systems (STS) [20]. However, the technical uncertainties

F. Bekius (✉)
Radboud University Nijmegen, Nijmegen, The Netherlands
e-mail: f.bekius@fm.ru.nl

S. Meijer
Delft University of Technology, Delft, The Netherlands

KTH Royal Institute of Technology, Stockholm, Sweden
e-mail: sebastiaan.meijer@sth.kth.se

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_7

cover just one part of the complexity. The behavior of, and interaction between, actors involved in the process is crucial for how the process evolves [8].

There is thus a need to specify the mechanisms and patterns actors perform in order to simulate the behavior of, and interactions between, actors involved. Game theoretical models describe these interactions and are used for social simulation [19–21, 32, 49]. Choosing the right mechanism for the situation at hand is crucial for simulation of real-world systems. However, selecting the right game concept is not trivial [6, 16].

In this short paper, we present a taxonomy of game theoretical concepts. The criteria used to design the taxonomy are important for selection of the *right* game concept given a real-world decision-making process. We choose a set of game concepts from this taxonomy that cover a wide range of situations to be used for social simulation of real-world systems. Thereby, we aim to enable social simulation designers by assisting decision-making processes.

7.2 Taxonomy of Game Concepts

The process of creating a taxonomy of game concepts and thereafter selecting a set of game concepts follows multiple steps. First, we list game concepts. Second, we define selection criteria and create a taxonomy of game concepts. Finally, we select a set of game concepts.

7.2.1 List of Game Concepts

First, we list a large number of game concepts by investigating game theory [31, 33, 40, 51] and public administration literature on complex networks [9, 11, 24–26, 48]. The list of game concepts can be found in Table 7.2 in the Appendix. Some game concepts are mathematically defined, such as the well-known prisoner's dilemma, while others have only been observed empirically. The multi-issue game is an example of a concept used in a rather descriptive, non-formal way [9]. Therefore, the characteristics of the game concepts vary between being empirically substantiated and mathematically proven.

The list of game concepts contains several concepts that resemble one another; this means, in game theoretical terms, they belong to the same class of games. We cluster the game concepts according to classes of games. In the game concept list, we indicate the main game class they belong to. However, since game concepts can belong to multiple classes, this does not help much with selecting game concepts. For example, the volunteers dilemma [13] belongs to the class of social dilemmas, symmetric games, and n-player games. Hence, we need to define appropriate criteria.

7.2.2 Criteria

The choice of criteria for categorizing game concepts is crucial for the selection of the *right* game concept. Taking into account our aim to simulate social interactions in real-world decision-making processes, some game theoretical classes are irrelevant. Distinguishing between complete and incomplete information games is game theoretically interesting, but when investigating real-world processes, a complete information game[1] almost never occurs since knowledge is not shared with everyone, and even if this is done, it is unlikely one knows exactly what is on the agenda of the other actors involved.

Therefore, we use criteria originating from theory on real-world complex decision-making processes and management of complex networks [9, 26, 48]. In these processes, multiple actors are involved, and the actors form a network of interdependencies. Within these networks hierarchical relations might exist; usually this is between two actors. Reaching a collective decision is the aim of the process. However, individual strategic behavior plays a role as well. Moreover, the decision-making process is dynamic. Some of these criteria overlap with the game theoretical classes, and this results in the following criteria for our taxonomy:

– Number of actors: two actors versus multiple (n) actors
– Relations and institutional structure: network versus hierarchy
– Strategy: collective decision-making versus individual strategic behavior
– Process and context: dynamic versus static

The taxonomy divides the list of game concepts in 16 groups; see for a graphical representation Fig. 7.1. An overview of the game concepts per group is presented in Table 7.1.

7.2.3 Selection of Game Concepts

From the game taxonomy, we select 10 game concepts that represent a wide range of interactions between actors in real-world complex decision-making processes. Hence, we are mainly interested in game concepts that describe multiple actors in a network structure who have to make a collective decision in a dynamic environment (group 9). However, we know that strategic behavior exists in the multi-actor setting (groups 11 and 12) and hierarchical relations as well (groups 13 and 15). On the other hand, interactions between two actors, potentially in a hierarchical relation, also impact the decision-making process (groups 1 to 8). In short, the selected game concepts should together span a large number of different groups of the taxonomy. The selected game concepts (bold concepts in Table 7.1) have different

[1] Knowledge about the game and the players in the game is common knowledge [40].

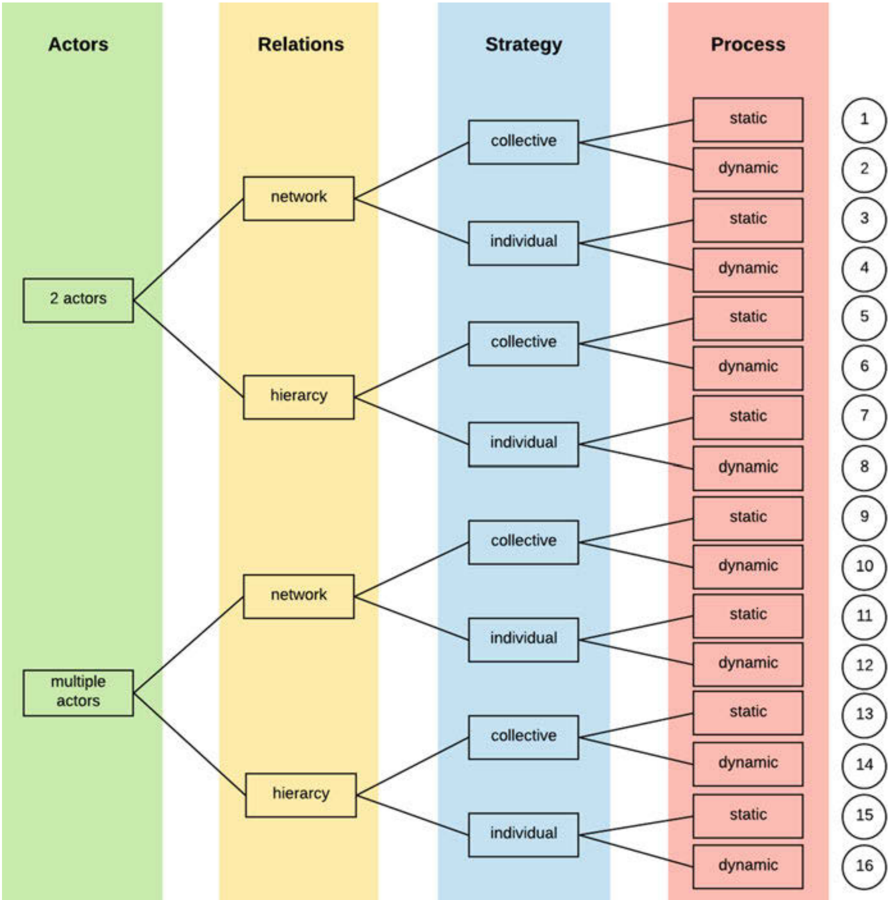


Fig. 7.1 Game taxonomy

characteristics. To identify the right game concept for the situation at hand, the elements of the situation need to match the characteristics of the game concept.

As a first step literature research is performed on the selected game concepts and key terms are defined [3, 4, 9, 12–14, 18, 27, 31, 33, 36, 38–40, 43, 45, 46, 51]. The process is terminated when two new papers did not add any new key terms. Subsequently, the key terms are clustered, and we describe the essence, context, process, results, and risks, i.e., the characteristics, for each game concept. To select a game concept for a given situation, we should be able to distinguish between them, and thus, we are interested in the distinguishable characteristics.

Taking this into account, we reconsider the selection of game concepts: camel nose [50] and allocation games [18] have characteristics that do not distinguish them from other game concepts. This means that given a certain situation in a decision-making process, it will be difficult to make them explicit, and therefore we decided

Table 7.1 Taxonomy of game concepts

(1)	Sender-receiver, grab the dollar, peel and pulp, **cascade game**, two-player unanimity, matching*
(2)	Stag hunt, **battle of the sexes**, chicken game, matching pennies, deadlock
(3)	Peel and pulp, **camel nose**, centipede game, peace war, war of attrition, Blotto games, Cournot games, cake division
(4)	**battle of the sexes**, stag hunt, prisoner's dilemma, traveller's dilemma, chicken game, matching pennies
(5)	**Two-level game**, **cascade game**, matching*
(6)	–
(7)	**Principal-agent game**, screening game, **camel nose**, inspection game
(8)	**Principal-agent game**, ultimatum game, trust game, dictator game, screening game, Stackelberg game
(9)	**Multi-issue game**, **allocation game**, **cascade game**, congestion game, tragedy of the commons, coalition games*, cooperative games*, signalling games*
(10)	El farol bar, voting games*
(11)	**Multi-issue game**, **allocation game**, **camel nose**, tragedy of the commons, Cournot games, signalling games*, auctions*
(12)	**Volunteers dilemma**; **diners dilemma**; public goods game; divide the dollars; el farol bar; rock, paper, scissors; voting games*
(13)	**Cascade game**, **two-level game**, signalling games*, coalition games*
(14)	Voting games*
(15)	**Hub-spoke**, **camel nose**, multiple-**principal-agent game**, signalling games*, mechanism design*, auctions*
(16)	Multiple-**principal-agent game**, Stackelberg game, voting games*

game* means that the game represents a larger class of games

to not incorporate them in our final selection. The characteristics of Two-level game [38] and Cascade game [14] turn out to be similar, and therefore, we select one of the two game concepts.

Hence, the final selection contains seven different game concepts: battle of the sexes, principal-agent game, multi-issue game, cascade game, hub-spoke game, volunteer's dilemma, and diner's dilemma.

7.3 Discussion

The proposed taxonomy is an attempt to provide some rigor in the selection of game concepts for social simulation. Social simulation can be used for decision-making processes in different forms: design of the decision-making process, providing input for the process, and evaluation of alternative outcomes of the process.

Our selection is a mix of policy-rich game concepts, originating from literature on policy networks, and concepts which are game theoretically defined. For the concepts to be useful for social simulation, they need to be rich enough to describe the complexity of the real-world system under study, and, on the other hand, they need to be formally defined. The characteristics of our selection of game concepts vary because of their different origins. This makes it difficult to specify the main characteristics and to distinguish between game concepts. Especially, being able to distinguish between game concepts is very important for the concepts to be useful for social simulation. For example, the power relation between actors is specified by the principal-agent game, i.e., the principal has more power than the agent. In the volunteers dilemma, power relations between actors are not made specific, they are able to perform the same actions, and thus we assume that actors have similar power positions. However, in real-world decision-making processes, we observe that the relation between actors in terms of power is of influence on how the volunteers dilemma evolves and who performs which actions. It is thus too simplistic to assume that power relations are not of influence when they are not specified. Therefore, additional empirical research is necessary to enrich the game theoretical concepts, such as the volunteers dilemma, while the policy-rich game concepts, such as the multi-issue game, require formalization.

7.4 Future Work

In the discussion, we mention two directions for future work:

(i) Formalization of game concepts originating from public administration. We are currently working on a formalization of the multi-issue game using computational social choice theory [7]. We use empirical data and observations from real-world decision-making processes, in which the multi-issue game is present, to decide on the variables to include and assumptions to be made.

(ii) Enriching the game theoretical concepts. We use several empirical case studies to identify the game concepts. This contributes to defining scenarios, risks, and opportunities of the game concepts and leads to more generic patterns of game concepts.

Concepts that are both formally defined, and rich in explaining different scenarios, contribute to enriching social simulation with formal game elements.

Another question is: how to select the *right* game concept, from the selection of seven game concepts, to enable decision-makers involved in complex decision-making-processes? We design an instrument that identifies game concepts. The instrument consists of questions regarding the decision-making process. These questions are based on the main characteristics of the game concepts. Arrows between the questions lead to game concepts. The instrument will be verified with experts in several rounds and, thereafter, validated in two ways. First, we test with hypothetical scenarios, based upon real-world complex decision-making processes. The research question to be answered is: does the instrument select the right game concept given a predefined scenario? Second, we test with real-world cases and decision-makers from the Dutch railway sector. The research question to be answered is: to which extend is the identification of game concepts useful for an organization? In a workshop setting, participants will identify game concepts and discuss the implications for the decision-making process.

Acknowledgements This research is funded, through the Railway Gaming Suite 2 program, by ProRail (the Dutch Railway Infrastructure Manager) and Delft University of Technology.

Appendix

Table 7.2 List of game concepts

Game concept	Game class	Reference	Game concept	Game class	Reference
Volunteer's dilemma	Social dilemma	[13]	Beer-Quiche game	Signalling game	[33]
Diner's dilemma	Social dilemma	[17]	Buyer-seller game	Signalling game	[33]
Public goods game	Social dilemma	[44]	Contract-signing game	Signalling game	[31]
Divide the dollars	Social dilemma	[33, 47]	Sender-receiver game	Signalling game	[33]
Blotto games	Social dilemma	[40]	Reputation game	Signalling game	[31]
Tragedy of the commons	Social dilemma	[52]	Selten's game	Signalling game	[40]
Sanitarian's dilemma	Social dilemma	[52]	Two-player unanimity	Coalition game	[33]
Battle of the Sexes	Coordination game	[40]	Landowner-worker game	Coalition game	[35]
Prisoner's dilemma	Coordination game	[40]	Three-player majority	Coalition game	[33]
Chicken game	Coordination game	[40]	Unanimity game	Coalition game	[36]
Stag hunt	Coordination game	[40]	Multi-issue game	Dynamic game	[9]
Traveller's dilemma	Coordination game	[31]	Cascade game	Dynamic game	[14]
Matching pennies	Coordination game	[40]	Hub-spoke game	Dynamic game	[1, 15]
El Farol Bar	Coordination game	[19, 28]	Peel and pulp	Dynamic game	[10]
Peace war game	Coordination game	[34]	Camel nose	Dynamic game	[50]
Deadlock	Coordination game	[31]	Centipede game	Dynamic game	[5]
Rock, paper, scissors	Coordination game	[28]	Inspection game	Dynamic game	[37]
Principal-agent game	Principal-agent game	[40]	Two-level game	Dynamic game	[38]
Multiple-principal-agent game	Principal-agent game	[27]	Allocation game	Resource allocation game	[12]
Dictator game	Principal-agent game	[27]	Cake division	Resource allocation game	[7]
Ultimatum game	Principal-agent game	[40]	War of attrition	Timing game	[35]
Trust game	Principal-agent game	[31]	Grab the dollar	Timing game	[40]
Screening game	Principal-agent game	[35]	Revelation principle	Mechanism design	[33, 40]
Marriage problem	Matching	[42]	Umbrella game	1-player game	[52]
School selection	Matching	[42]	Minority game	Congestion game	[41]
Kidney exchange	Matching	[42]	Voting game	Voting game	[7]
Nontransferable utility	Cooperative game	[36]	Cournot game	Cournot games	[47]
Transferable utility	Cooperative game	[36]	Stackelberg game	Stackelberg games	[33]
Hedonic games	Cooperative game	[2]	Dollar auction	Auctions	[40]

References

1. Adler, N.: Hub-spoke network choice under competition with an application to Western Europe. Transp. Sci. **39**(1), 58–72 (2005)
2. Airiau, S.: Cooperative games and multiagent systems. Knowl. Eng. Rev. **28**(4), 381–424 (2013)
3. Anderson, L.R., Holt, C.A.: Classroom games: information cascades. J. Econ. Perspect. **10**(4), 187–193 (1996)
4. Archetti, M.: Cooperation as a volunteer's dilemma and the strategy of conflict in public goods games. J. Evol. Biol. **22**, 2192–2200 (2009)
5. Aumann, R.: Backward induction and common knowledge of rationality. Games Econom. Behav. **8**(1), 6–19 (1995)
6. Barreteau, O., Le Page, C., Perez, P.: Contribution of simulation and gaming to natural resource management issues: an introduction. Simul. Gaming **38**(2), 185–194 (2007)
7. Brandt, F., Conitzer, V., Endriss, U., Lang, J., Procaccia, A.D.: Handbook of computational social choice, 1st edn. Cambridge University Press, New York (2016)
8. de Bruijn, H., Herder, P.M.: System and actor perspectives on sociotechnical systems. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **39**(5), 981–992 (2009)
9. de Bruijn, H., Ten Heuvelhof, E.: Management in Networks: On Multi-Actor Decision Making. Taylor & Francis Ltd, Routledge (2018)
10. de Bruijn, H., ten Heuvelhof, E., in 't Veld, R.: Process Management, Why Project Management Fails in Complex Decision Making Processes. Springer, Berlin/Heidelberg (2010)
11. Bueren, E.M., Klijn, E.H., Koppenjan, J.F.M.: Dealing with wicked problems in networks: analyzing and environmental debate from a network perspective. J. Public Adm. Res. Theory **13**(2), 193–212 (2003)
12. Chevaleyre, Y., Dunne, P.E., Endriss, U., Lang, J., Lemaître, M., Maudet, N., Padget, J., Phelps, S., Rodríguez-Aguilar, J.A., Sousa, P.: Issues in multiagent resource allocation. Informatica **30**, 3–31 (2006)
13. Diekmann, A.: Volunteer's dilemma. J. Confl. Resolut. **29**(4), 605–610 (1985)
14. Easly, D., Kleinberg, J.: Games. In: Easley, D., Kleinberg, J. (eds.) Networks, Crowds, And Markets: Reasoning About A Highly Connected World, chap 6, pp. 155–208. Cambridge University Press, Cambridge (2010)
15. Elrod, J.K., Fortenberry, J.L.: The hub-and-spoke organization design: an avenue for serving patients well. BMC Health Serv. Res. **17**(1), 457 (2017)
16. Feld, S.L.: Simulation games in theory development. Sociol. Forum **12**(1), 103–115 (1997)
17. Gneezy, U., Haruvy, E., Yafe, H.: The inefficiency of splitting the bill. The Economic Journal, **114**(4), 265–280 (2004)
18. Goeree, J.K., Holt, C.A.: An explanation of anomalous behaviour in binary-choice games: entry, voting, public goods and the volunteers' dilemma. In: Virginia Economics Online Papers. University of Virginia, Department of Economics (2000)
19. Goeree, J.K., Holt, C.A.: Stochastic game theory: for playing games, not just for doing theory. Proc. Natl. Acad. Sci. **96**(19), 10564–10567 (1999)
20. Helbing, D., Balietti, S.: From social simulation to integrative system design. Eur. Phys. J. Spec. Top. **195**(1), 69–100 (2011)
21. Helbing, D.: A mathematical model for the behavior of individuals in a social field. J. Math. Sociol. **19**(3), 189–219 (1994)
22. Holland, J.H.: Complex adaptive systems. Daedalus **121**(1), 17–30 (1992)
23. Holland, J.H.: Studying complex adaptive systems. J. Syst. Sci. Complex. **19**(1), 1–8 (2006)
24. Kickert, W.J.M., Klijn, E.H., Koppenjan, J.F.M.: Managing Complex Networks. SAGE Publications Ltd, London (1997)
25. Klijn, E.H., Teisman, G.R.: Strategies and games in networks. In: Kickert, W., Klijn, E.H., Koppenjan, J. (eds.) Managing Complex Networks, chap. 6, pp. 98–118. SAGE Publications Ltd, London, UK (1997)

26. Koppenjan, J., Klijn, E.H.: Managing Uncertainties in Networks. Taylor & Amp; Francis Ltd, London, UK (2004)
27. Laffont, J.J., Martimort, D.: The Theory of Incentives: The Principal-Agent Model. Princeton University Press, Princeton and Oxford (2002)
28. Leyton-Brown, K., Shoham, Y.: Essentials of Game Theory: A Concise, Multidisciplinary Introduction. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool Publishers, Oregon State University (2008)
29. Meijer, S.A.: Gaming simulations for railways: lessons learned from modeling six games for the Dutch infrastructure management. In: Perpinya, X. (ed.) Infrastructure Design, Signaling and Security in Railway, pp. 275–294. Intech, Rijeka (2012)
30. Miller, J.H., Page, S.E.: Complex Adaptive Systems: An Introduction to Computational Models of Social Life. Princeton University Press, Princeton (2007)
31. Morrow, J.D.: Game Theory for Political Scientists. Princeton University Press, Princeton (1994)
32. Moss, S.: Game Theory: Limitations and an alternative. J. Artif. Soc. Soc. Simul. **4**(2) (2001)
33. Myerson, R.B.: Game Theory: Analysis of Conflict. Harvard University Press, Cambridge (1997)
34. O'Neill, B.: Game theory models of peace and war. In: Aumann, R.J., Peyton, H. (eds.) Young Handbook of Game Theory with Economic Applications, vol. 2, pp. 995–1053. Elsevier, Amsterdam (1994)
35. Osborne, M.: An Introduction to Game Theory. Oxford University Press, New Delhi (2003)
36. Osborne, M., Rubinstein, A.: A Course in Game Theory. The MIT Press, Cambridge (1994)
37. Owen, G.: Game Theory. Academic, Inc., New York (1982)
38. Putnam, R.D.: Diplomacy and domestic politics: the logic of two-level games. Int. Org. **42**(3), 427–460 (1988)
39. Ragin, C.C., Becker, H.S.: What is a Case? Exploring the Foundations of Social Inquiry. Cambridge University Press, New York (1992)
40. Rasmusen, E.: Games and Information: An Introduction to Game Theory. Blackwell Publishing, Malden (2007)
41. Rosenthal, R.W.: A class of games possessing pure-strategy Nash equilibria. Int. J. Game Theory **2**(1), 65–67 (1973)
42. Roth, A.E., Sotomayor, M.: Two-sided matching. In: Aumann, R.J., Hart, S. (eds.) Handbook of Game Theory with Economic Applications, vol. 1, pp. 485–541. Elsevier, Amsterdam (1992)
43. Scharpf, F.W.: Games Real Actors Play, Actor-Centered Institutionalism in Policy Research. Westview Press, Oxford (1997)
44. Schelling, T.: The Strategy of Conflict. Harvard University Press, Cambridge (1960)
45. Sebenius, J.K.: Negotiation arithmetic: adding and subtracting issues and parties. Int. Org. **37**(2), 281–316 (1983)
46. Stauvermann, P.J.: Application of principal-agent models in science policy. In: Comparative Perspectives on Scientific Expertise for Public Policy, Amsterdam, the Netherlands, pp. 1–13 (2004)
47. Straffin, P.D.: Game Theory and Strategy. Mathematical Association of America, Washington, DC (1993)
48. Teisman, G.R., Klijn, E.H.: Complexity theory and public management. Public Manag. Rev. **10**(3), 287–297 (2008)
49. Vollmer, H.: What kind of game is everyday interaction? Ration. Soc. **25**(3), 370–404 (2013)
50. Volokh, E.: The mechanisms of the slippery slope source. Harv. Law Rev. **116**(4), 1026–1137 (2003)
51. Von Neumann, J., Morgenstern, O.: Theory of Games and Economic Behavior. Wiley, New York (1953)
52. Webster, T.: Introduction to Game Theory in Business and Economics. M.E. Sharpe, Inc., New York (2009)

Chapter 8
Physician, Heal Thyself! The Prospects for Using ABM to Target Interventions to Raise ABM Engagement

Edmund Chattoe-Brown

Abstract This paper presents a preliminary "structural" model of a simplified educational system. This model is used to analyse different strategies for enhancing the engagement with ABM in the education system. Another novel feature of the model is that it considers both formal and informal channels of education (e.g. not just classroom teaching but also blogs). Subject to its assumptions, the ABM supports a pre-existing view that uptake may be enhanced by making teaching available earlier in the education process. Analysis also suggests that a combination of formal and informal learning may be more effect than the simple addition of each effect. The paper also considers the promotional and intellectual benefits of modelling educational systems.

Keywords Education systems · Curriculum planning · Innovation diffusion · Formal and informal learning · Agent-based modelling · Research methodology

8.1 Introduction

This abstract presents an argument and outline for an ABM that can underpin effective discussion of, and intervention in, the educational system with the somewhat self-serving aim of promoting the uptake and sustainability of ABM as a research method. The issue is that the education system would seem to be a paradigm case of a complex system where the consequences of plausible interventions may not be at all those that are anticipated. To take a homely example that has already been discussed in the ABM community, it seems to make sense to try and "catch" potential recruits to ABM earlier. (Training only at the level of master's – so-called postgraduate taught or PGT – is already talking to a very small pool, even if a very competent one.) But what is behind such a view and is the intuition actually

E. Chattoe-Brown (✉)
School of Media, Communication and Sociology, University of Leicester, Leicester, UK
e-mail: ecb18@le.ac.uk

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_8

correct? Mere familiarity with an approach at an early stage may shape choice, but how big is this effect if motivation and performance must also be sustained? Is it no good reaching students in school if we cannot also reach them as undergraduates, or are we still “priming” them efficiently for greater participation in relevant master’s degrees? Is the current “thinly spread” training offered by particular departments effectively negated by the fact that almost no student will choose a degree based on a single taught component? Should this phenomenon encourage us to think about collective action failures (within a competitive and neoliberal education system) and alternative collective modes of teaching delivery? Given that we have very limited resources (taking our competence and enthusiasm as given), how can we make best use of these? In almost any other sphere, we would be advising the use of an ABM in this situation. Do we not believe our own advertising? (This self-referential situation has other appealing features. If we cannot build a model to help ourselves, how can we commend this to others? Should we not be able to build a really impressive model together in our own interests that can then serve as an exemplar of what can be achieved?) The next section discusses what such a model would need to do. The following section discusses some extremely tentative results to illustrate the issues involved. The final section concludes.

8.2 A Model Proposal

This model builds on earlier work exploring the transmission and sustainability of expertise [1]. The basic ideas are simple. The capability to achieve high levels of expertise is rare. It probably requires a mixture of unusual innate ability and teaching that is itself expert to create more expertise. How does the configuration of the education system sustain (or perhaps even increase) levels of expertise? Behind these questions are many interesting issues. How can we represent (or create) an education system that teaches people to surpass their teachers? To what extent is the sustainability of education dependent on nurturing exceptional individuals in a distinctive way? Conversely, do we have to worry that education systems can, under certain circumstances, steadily degrade expertise while allowing that fact to remain invisible?

As ever, for a first model, much simplification is needed (though the final section discusses where this is likely to be most problematic and what we can do about it). In the model presented here, agents are born with three “competences” (call them reading, riting and rithmetic) and a potential competence in ABM. There are two sectors of education at all levels, an “elite” one (which selects the best of the students into it up to its capacity) and a “residual” one that takes everybody else. There are six levels of education (primary, secondary, college, undergraduate – hereafter UG, master’s – hereafter PGT and PhD) of which only the first two are compulsory. The first three levels of education have the same “capacity” (which is sufficient for all children) of 20 in the elite institution and 20 in the residual institution. At UG and PGT level, the capacity is reduced to 10 and 10 and at PhD level to 5 and 5.

Broadly basing this on the UK system and simplifying extremely (so, e.g. there is no "Common Entrance" system to certain private schools at age 13), all children go to school at 5 and spend 5 years at primary school, 6 years at secondary school, 2 years at college, 3 years at UG and 1 year at PGT. (The ABM does not explicitly deal with what happens at the PhD level since, at present, it is only interested in comparing the outcomes of different arbitrary teaching arrangements although, ultimately, the supply of PhD students "loops back" to the next generation of university teachers.) At present, learning simply involves incrementing ability levels in each competence probabilistically. (Motivation, class composition and teaching effects will be added shortly.) Children arrive at primary school with a distribution of competences that they add to (based on "capability" – an attribute somewhat like IQ) at different rates. This means that by the time they come to be selected for secondary school, some are performing better than others (most, but not all of these, will probably be from the elite primary school). The best students are then selected into the elite secondary school and so on up the "educational ladder". The selection criterion is somewhat artificial in that it relies on "godlike" knowledge of all the student abilities to set just that entrance requirement that ensures that the elite institution is filled to capacity. (Real institutions would have to be adaptive based on whether they found themselves to be oversubscribed or undersubscribed.) Note that this means that the entrance qualification will be higher once the capacity of the elite institution becomes more restricted. (At higher levels, not all students can access UG, let alone elite UG.) It is assumed that there is a government grant scheme (again on the old English model), so anyone who is capable of going to university can afford to do so, and thus all students apply to the best institution that will admit them.

The final element of the basic model is that not all learning opportunities (particularly for unusual topics like ABM) come from the formal education system. There may also be popular books, magazines and newspapers, blogs and web sites and so on. This extramural learning process is represented by patches (the ABM is written in NetLogo and is available on request from the author), which have an "access level" (the competence a student must already have achieved to benefit from them) and a "delivery level" (the added competence that the resource can deliver after it has been accessed). This means that, provided a student has the relevant competence level, it can be "advanced" by self-study.

This model is extremely simple, but being the first of its kind as far as I can tell (though I am continuing to investigate), this simplicity is legitimate. Furthermore, it is still sufficiently complex for the results' section to consider two interesting issues in "resource use" for the propagation of ABM already suggested above. The first is how much effect the level at which ABM teaching is first introduced (we assume it enters as part of one of the other competences at a particular level) has on the level of competence in the population. The second is what effect the balance between devoting resources to formal teaching (UG modules in ABM) and "promotional activity" (like blogs or self-teaching resources) might be.

Like a lot of such initial models, part of the aim is to start a dialogue, not only to "capture intuitions" (and their diversity) from the modelling community but to think about the extent to which such models are compatible with data for calibration and validation (either actual or potential). We have a *lot* of data about school achievement in general terms, but how can it be linked to the much smaller scale data about, for example, the academic careers of successful (or just persistent) academics in ABM? Do those with careers in ABM owe this to higher ability, concentration of teaching in elite institutions or other things? Should we aim to "democratise" ABM training rather than simply offer more of it? However speculative this ABM is, it is the only one I have been able to find that attempts to deal with the education system as a whole. (And, of course, the closing of the loop is how those at the "end" of education do – or don't – become teachers.)

8.3 Some "Results" for Discussion

These results are very tentative. I am not even yet completely assured that they might not involve programming bugs. Nonetheless they are sufficiently interesting to present with that caveat.

Figure 8.1 shows (average of ten simulation runs for each condition after 11,000 ticks – by which point the whole education system has reached a somewhat steady state) the number of agents who have acquired an arbitrary level of competence ("2 points") under different "educational systems". In one education system, formal ABM teaching is only introduced at level 5 (PGT), while in another it is introduced at level 4 (UG). In addition, some systems allow students to learn external to formal teaching, while others do not.

The results are of some interest though they must be treated with caution even apart from the possibility of bugs and the arbitrary assumptions. Firstly, it is pretty clear that launching formal teaching earlier has a huge effect. In a way, this is not surprising. However hard it is to learn something, people will probably do so if you give them long enough. What is more interesting, however, is the effect that "informal" learning has. Taken alone it has a respectable effect, but this average conceals a large number of runs where nobody learns anything and a small number where people learn a lot. This situation is analogous to an issue that arises with Granovetter's threshold model [2]. Because the requirements and outcomes of the informal learning patches are randomly generated, it is easy for runs to exist where students either cannot "get started" (there are no resources that work from zero competence) or cannot get past a certain level (there are no resources "bridging a gap" in the competence "ladder"). This is an abstract result but has a correspondence to a real result. In the absence of much formal teaching in ABM, how confident are we that students can "self-study" to any arbitrary level of competence?

What is most interesting is what happens when we combine the effects. Social learning added to initial formal teaching at level 5 has a noticeable effect (though it is small in absolute terms because it builds on such a small base), and the effect is even more proportionately noticeable when teaching starts from level 4. The number of students who reach a given level of competence under the combined effects is also noticeably greater than the number that succeeds with each effect alone. (However,

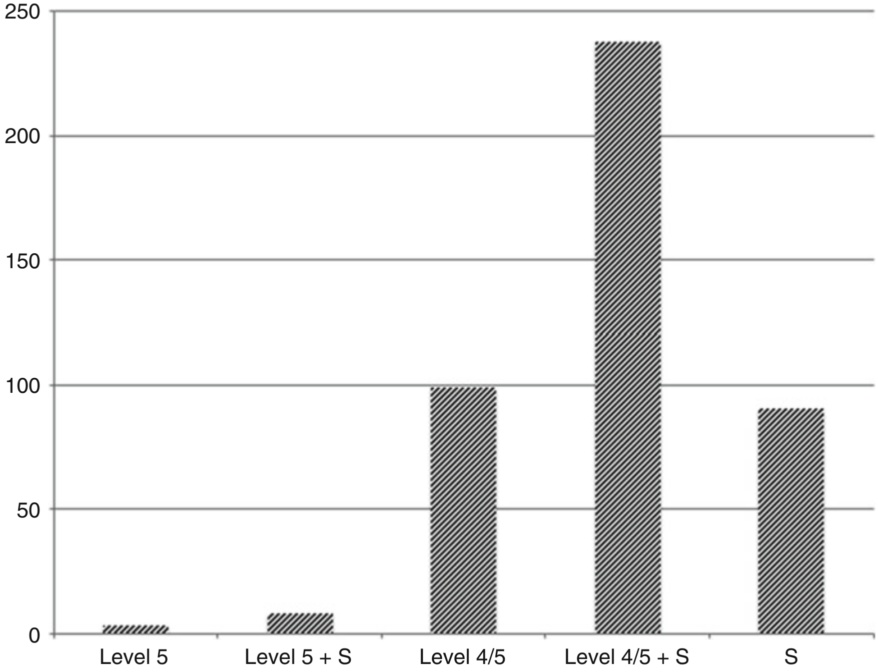


Fig. 8.1 The effect of different learning systems on the number of students acquiring an arbitrary level of ABM competence. Level 5 means "formal" ABM teaching only at PGT (master's) level. Level 4/5 means ABM teaching also at UG level. S means there is (or there is also) learning outside the classroom (i.e. self-study texts, blogging, newspaper articles and so on)

more simulation runs need to be done because, unless there is a bug, it is not clear how "level 5 + S" can be worse than "S alone". My best guess here is that the problem is the very uneven distribution of runs where informal learning achieves nothing at all.)

How interesting and surprising this result is needs further thought, but it does suggest a way of looking at education that could have more general heuristic (and perhaps even policy) value. Education is like a ladder where the job of each set of teachers is to "raise up" their students from the point where they take possession of them to the point where they can hand them on. If, for any reason, this ladder becomes interrupted, all higher stages are potentially compromised, and the more broken the ladder is, the worse the problem gets at higher levels. However, in opposition to this negative idea, the balance of formal teaching and informal resources may work better than either approach alone. The informal resources may carry at least the keenest students' past gaps in formal provision, and formal teaching may be a good way to ensure that informal resources are "joined up" where students can't proceed alone. But thinking of education as a collective system (with a potential for collective failure) constitutes an important basis for better understanding such challenges (as ABM enthusiasts often argue to others!)

8.4 Conclusions

In a way, despite the novel application area and the slightly unusual abstraction, this ABM is the start of a perfectly standard modelling agenda. It is designed to start by building on existing models (though there don't seem to be any), synthesising competing theories of the phenomenon (although again these seem to be "popular" to the extent that they exist at all: Is there really a "theory" of "dumbing down?") and "putting in order" existing data that might be used for calibration, validation or "structural validity" (the idea that the model doesn't exclude any process or mechanism for which there seems to be evidence of impact, so there are currently no "class effects" in the model. Is this legitimate as an approximation to what we actually know about class effects?) In fact, despite the abstract nature of the model, the data is perhaps not as abstruse as it might appear. We could start with biographical interviews with ABM practitioners and also interview a sample of wider populations to find out who has heard of the method, been interested in it, tried to use it, been discouraged and so on.

In a way, what is interesting is the opportunity this project allows the ABM community (or at least those interested enough to read this paper) to "put its money where its mouth is" regarding the value and feasibility of building ABM to assist in practical problem-solving. We *ought* to be able to get this to work for us in our own interests. If we can't, then we should be worried. (And this includes the worry that, in practice, many people find evidence-based policy too much like hard work and would rather just pontificate on the basis of personal opinion!)

However, having said that this model has a number of abstractions and simplifications, it is not so unusual in approach as to be potentially distortionary. There is no reason why later versions of the model should not relax the assumption that students always want to learn (and that their motivation is not enhanced by success and diminished by failure). Similarly, although the model currently contains teachers, they do not really do anything, but it would not be at all different to have them deliberately arrange teaching to take account of the ability of their current students and/or devote some time to getting specific students "unstuck" (able to progress for themselves) or "remotivated" (feeling it is more worth continuing the attempt to learn).

As with any practical modelling challenge, new domains throw up interesting new issues. In this case, while structural ABMs exist, they tend to be more deterministic than the aspiring model presented here. What happens when you have *both* different possible school structures and individual teacher agency? How much difference do structures actually make? (Education research has a number of counter-intuitive results that may be susceptible to ABM, for example, that class size has a very limited effect of performance.) Furthermore, rather few ABMs consider the role of material artefacts (like books) in the knowledge process. In this model, these serve as "fixed points" that can operate to counter-balance the "socially constructed" nature of the teaching process, dealing as it has to with the "typical student" at any given level in any given institution. (The model has not yet dealt with the interesting issue that different artefacts are “available” over different time scales from the permanence of books in public libraries to the very transient nature of blog posts. But the existing ABM framework is capable of supporting this insight with very little additional programming.)

Finally, the model may have the potential for wider use. Boosting ABM is not the only reason for studying the complexities of whole education systems and possible policy interventions! At the very least, other minority subjects (like Latin) may also have questions to ask about strategies to improve their sustainability. And, of course, this approach (or a more developed version of it) could have been a godsend to the debate between “comprehensive” education and differing routes through the school system with a more academic or vocational emphasis.

References

1. Chattoe-Brown, E.: Simulating “stage” theories: an exploration with applications. The Sixth Conference of the European Social Simulation Association (ESSA 2009), Guildford, UK, 14–18 September, paper presentation
2. Granovetter, M.: Threshold models of collective behavior. Am. J. Sociol. **83**, 1420–1443 (1978)

Chapter 9
So You Got Two Ologies? The Challenge of Empirically Modelling Medical Prescribing Behaviour and Its Effect on Antimicrobial Resistance as a Case Study

Edmund Chattoe-Brown, Andrew M. Colman, David R. Jenkins, Eva M. Krockow, Shaheen Mehtar, Nelun Perera, and Carolyn Tarrant

Abstract This paper reports the early stages of a funded project "Antimicrobial Resistance as a Social Dilemma: Approaches to Reducing Broad-Spectrum Antibiotic Use in Acute Medical Patients Internationally" to use ABM in understanding how the prescribing behaviour of professionals in "medical institutions" may affect the prevalence of antimicrobial resistance (hereafter AMR), a situation in which standard treatments for infections fail. The first section explains the policy challenge of AMR and the research project strategy. The second section considers distinctive challenges raised by this kind of policy modelling. The third section presents preliminary results and the final section concludes.

Keywords Policy modelling · Antimicrobial resistance · Antibiotic prescribing · Coupled models · Health policy · Agent-based modelling · Research methodology · Collective action problems

E. Chattoe-Brown (✉)
School of Media, Communication and Sociology, University of Leicester, Leicester, UK
e-mail: ecb18@le.ac.uk

A. M. Colman
Department of Neuroscience, Psychology and Behaviour, University of Leicester, Leicester, UK

D. R. Jenkins · N. Perera
Department of Clinical Microbiology, University Hospitals of Leicester NHS Trust, Leicester, UK

E. M. Krockow · C. Tarrant
Department of Health Sciences, University of Leicester, Leicester, UK

S. Mehtar
Tygerberg Academic Hospital and Faculty of Health Sciences, Stellenbosch University, Cape Town, South Africa

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_9

9.1 The Policy Context

AMR is a current (and potentially catastrophic future) global problem in healthcare (http://www.bbc.co.uk/news/health-21702647). Formerly, the development of new antibiotics led medical professionals to believe that new treatment possibilities would always be available, but more recently, there has been a "drying up" of new medicines and a decline in the effectiveness of current ones. In a nutshell, AMR arises when bacteria are exposed to medicines that lack sufficient strength to kill them outright. (This may happen for many reasons: failure to complete the course of treatment, mis-prescription, substandard drugs, the presence of asymptomatic conditions that are not intentionally being treated and so on.) The project on which this research is based is specifically organised around the problem that prescribing behaviour may constitute a social dilemma [1]. Ideally, medical professionals would use the results of laboratory tests to accurately identify bacteria and thus prescribe the most effective and least "broad-spectrum" antibiotic. (Narrow-spectrum antibiotics are effective against one or a few conditions, while broad-spectrum ones have wide effectiveness.) In practice, medical good practice and the speed with which some conditions (like sepsis) develop make this strategy unfeasible (particularly where laboratory tests may be delayed or unreliable) with the result that medical professionals frequently begin by prescribing broad-spectrum antibiotics. These are rather likely to be effective against whatever ails the patient but at the cost of potentially creating resistance in other bacteria at the same time (which can then reproduce and spread in the population by a variety of mechanisms).

How do we manage a system in which the individual benefits to the particular patient (not dying) may be in conflict with the benefits to future patients (having effective treatments so *they* don't die?)? What would it be like to live in a world where currently routine surgery (for example, caesarian sections and appendectomies) was too risky because typical concomitant infections had become untreatable? There is also a wider context in which the use of antibiotics in animals is not limited to medical need – regular doses promote growth even in healthy animals – and not regulated with AMR in mind and where, in many countries, either deliberately or through inadequate enforcement, strong antibiotics can be bought "over the counter" without the intervention of relevant medical expertise serving a stewardship role [2]. This is another example of a global health problem. It doesn't matter where AMR arises – and it is rather likely to arise at the "weakest links" in the global health system. Once it has arisen, it *will* spread.

This paper reflects on a variety of challenges that ABM faces in making a worthwhile contribution to this problem in the hope that these may serve as a case study of interest for other such projects.

9.2 The Policy Challenges

Part of the reason for presenting this model (even in its early stages) is that it throws up a number of interesting challenges of wider potential relevance. These can be divided into three broad classes.

The first class (which inspired the title of the paper) is the need to build a model which interfaces directly both with microbiology and with epidemiology (as well as the psychology and sociology that bear on medical decision-making and the institutional structure of medical treatment). In order to understand the problem of AMR, it is not possible to "abstract out" either from the way that illnesses develop and present to the patient and the medical professional or from the way that these illnesses spread from person to person in the "wider world". This immediately presents a subsidiary problem (particularly for illness presentation) that might be called "model scaling". The initial ABM of decision-making (mainly designed as a basis for discussion by various experts and research collaborators – and described in much more detail in [3]) is neither large nor hugely complicated, but it is not at all clear to what extent it is possible both to abstract realistically from microbiology (such that the policies developed using the ABM stand some chance of actually being useful in the real world) and to ensure that the microbiological element of the model doesn't become unfeasibly large in its own right (but see [4]). There are two specific aspects of microbiology that might give rise to this challenge. The first is the process of mutation and transmission of resistance. In this process, completely new mechanisms of resistance (either for specific antibiotics or for a range of treatments operating in a similar way) can arise and, once they have arisen, can transfer from one kind of bacteria to another (resulting in the ultimate threat of multidrug-resistant bacteria). The second is the detail of the interplay between bacteria and antibiotics. For example, an antibiotic that is introduced into the body in sufficient doses to treat sepsis in an external wound effectively may then operate at a much lower dose in the gut and thus create potential for the development of AMR at that site (particularly since the gut is often a site of asymptomatic bacterial colonies). Furthermore, there is a crossover effect in that one thing that determines the more "traditional" (Darwinian) spread of AMR is whether or not particular mechanisms of resistance (or combinations of mechanisms) experience a significant "evolutionary survival penalty" in their reproductive cost. If the antibiotic is present and "challenging" the bacteria, it may be a resistant variant that comes to dominate the environment (with consequences for subsequent transmission), but if the antibiotic is absent, the non-resistant variant may be the one that has greater reproductive potential. It can be seen that if the ABM needs to represent individual bacteria (or even competing open-ended populations of variant bacteria) at different sites in the body, then the model needed as a "chassis" for understanding the impact of medical decision-making could become extremely challenging in its own right. (It would be frustrating to invest in such a model only to discover it had little effect on the policy outcomes, but in a complex system, such a discovery is perfectly possible.) The attempt to address this problem is ongoing but has proceeded by stages. The initial ABM (discussed below) was designed to be concise and "not incompatible" with relevant reading and expert judgment. However, while it was possible to devise a coherent and "structurally plausible" ABM (one that did not abstract out from any clearly relevant processes like prescribing decisions or social transmission of disease), it was also clear that the model was not congruent with certain aspects of the intended problem under study (in particular the distinctive functioning of broad- and narrow-spectrum antibiotics in prescribing and the exact nature of the resulting social dilemma). However, the basic model proved sufficient both to sharpen thinking about the nature of this challenge and devise effective questions that allowed the elicitation of a stylised microbiology from an expert microbiologist. (It is surprising how reticent published academic research on AMR is regarding how this is actually supposed to occur.) This stylised model is currently being finalised and incorporated into the ABM.

The second challenge (which might be seen as an outgrowth of the first) is to build and interpret a model in which the diverse perceptions of different types of agents shape the evolving system. There is an explicit microbiology that researchers can "access" via their privileged position with respect to the ABM, but the patient only knows that they "feel bad" and can articulate certain symptoms (like fever or diarrhoea). The medical professional can diagnose on the basis of those symptoms (though they may not be perfectly reliable indicators of the underlying condition and the patient may not be able to report them reliably) or can make use of laboratory testing (which "bypasses" the symptoms and attempts – but still somewhat fallibly – to identify the actual infecting agent). Back in the real world, notoriously, we only have data about patients who actually seek treatment (unless they die which produces *some* data but not necessarily what is needed), and even if we know what they are suffering from, we may not know with any assurance where they got it. (This is a problem besetting models that attempt to adjudicate the different "causes" of hospital infections such as different classes of medical workers and poor hospital hygiene; see, for example, [5].) By its nature, this problem cannot be "solved", but it is far better that the modeller be aware of it than not be. As suggested in the discussion of results below, the best that can be hoped for is a model that sensibly "interfaces" with the sorts of data that are actually likely to be available at different levels of description such that they can then be used for calibration and/or validation as appropriate. For example, we can ask patients how they feel directly. We can ask medical professionals to diagnose based on different kinds of symptoms (which may be "invented" to explore counterfactual decision-making). We can use laboratory research to check the connections between tests and symptoms: To what extent is diarrhoea a reliable indicator or a particular illness? To what extent can it be made a more reliable indicator by asking subsidiary questions (short of testing) about that symptom? It seems that (based on this example) many policy models may have to engage with the fact that different actors do not simply disagree about a shared state of the world (e.g. as with economic preferences and many analytical models) but actually see the world in fundamentally different ways. (For an example of subjective and "objective" data in the slightly different context of wellbeing, see [6].)

The final challenge (and one that is less specific to the AMR model but is nonetheless interesting) is to make effective use of different sources of data to end up with something that is actually suitable to make a policy contribution. One can of course build the model using expert opinion alone (and this may have rhetorical advantages in getting "implementation buy-in"), but a model that simply reproduces particular prejudices may do more harm than good (see [7] for a recent general introduction to this approach). How do we draw a boundary between scientific knowledge that can be "objectively" incorporated into the model (whether the experts are comfortable with it or not) and knowledge that is *only* effectively accessible via experts (and filtered through their potential prejudices)? Some relevant examples, showing the problematic link between funded research and practical science, have already been suggested. Firstly, it may be that any benefits of a project to improve prescribing for humans are "swamped" by the effects of animal use and/or free-market availability of antibiotics. Even if it becomes clear that market aspects are the problem, a lack of political will and the limited feasibility of effective regulation in some countries mean that it almost certainly will not be dealt with. Secondly, it is surprisingly hard to establish conclusively from the literature what the evidence base is for the claim that prescribing behaviour creates or speeds up the development of AMR. While it is fairly plausible (and supported by qualitative interviews with relevant medical professionals) that there is a potential social dilemma in play, one simple but fundamental reason why prescribing might not be susceptible to change is because medical professionals simply don't share the characterisation of the problem put forward by the researchers. Thus there is not only a technical problem of making the ABM both empirically convincing and suitable for exploring practical policy but a *contextual* problem of making sure it is solving the *right* problem and cannot simply be dismissed by experts as enshrining a particular arbitrary view that they do not share. We now turn to seeing how some of these issues play out in the current form of the model.

9.3 Some Preliminary Results

As already suggested, the aims of this model were, firstly, to produce something quickly that worked and could then be a basis for collaborator and expert discussion and, secondly, to create something approximately "structurally valid" for future development. By the latter, we mean that while particular elements of the model were simplified (e.g. medical professionals always have access to exactly the right antibiotic for any given illness), there were no aspects of the domain (at least based on our judgment of the literature and considerable expert discussion) that was deliberately excluded. The pursuit of this aim is less common than might be supposed. To take just one example, McPhee-Knowles [8] opens her discussion of a food safety ABM by discussing the implications of "the increasing complexity of the food supply chain", but her model contains no such supply chain, and even the shops and restaurants she does represent are not "correlated" in sharing suppliers (as might be expected if supply chains were in operation). The first ABM thus serves as a "chassis" for later versions (although, as previously suggested, the stylised microbiology of the initial version has already been found wanting and is in the process of being superseded.)

A good illustration of this "structural validity" approach is the treatment of medical professionals in the model. The research project deals with hospitals and, as such, will need to take account of institutional structures and things like the possibility of cross infection (see, for example, [9]). The simple model involves what are effectively "single-hander" GP practices so it retains the structural aspects of patients deciding they are ill and needing treatment (and then going to a medical professional and receiving it) but does not, at this stage, deal with the "bureaucratic detail" of that process (and to be fair, in the UK at least, much illness *is* dealt with effectively on the basis of single GP visits followed by dispensing via a chemist/pharmacist.) A similar principle of simplification without elimination was the aspiration of the model as a whole.

The model has three core elements: the treatment process, the illnesses and the social context. The latter is probably the simplest to describe. Agents move around at random (another simplification from moving around purposively that nonetheless retains the "structural importance" of movement and social contact) and may thus infect each other with any illnesses they have (probabilistically and on the assumption that all illnesses are based on mere contact rather than, say, sexual activity – again an intended attempt at "structural validity" in regard to transmission). If they decide they are ill, they will move immediately to a patch (the simulation is written in NetLogo, and the code is available from the corresponding author) that has a medical professional on it. (These are a small fraction of the total population.) The medical professional will look at their symptoms (there are no laboratory tests at present though they will be implemented very shortly) and then decide what treatment to prescribe. The patient will then, 100% reliably, participate in the treatment and will usually recover (for reasons that are just about to be explained).

As suggested above, in a sense, the illnesses need to be the most complicated part of the whole ABM. The mechanism of illness presentation is designed to be, in a stylised way, a reflection of how illness appears to work simultaneously from the perspective of the microbiologist, the patient and the medical professional. The key concept in the stylised representation is that of "illness level" (which might represent the quantity of bacteria in the body, for example, and thus be calibrated at least approximately). When first infected, a patient has a very low illness level and no symptoms (but they will be contagious). As the illness level rises, symptoms appear, at which point the patient is in a position to seek medical treatment. If they don't do so, or if the treatment doesn't work, their illness level continues to rise until they develop symptoms that either carry a risk of death (like convulsions) or directly result in death (like heart failure). Illness level is a simple "point total" that is incremented stochastically in each tick with two thresholds (for symptoms appearing and risk of death). Furthermore, illness level can be decremented either by "natural resistance" or by treatment. (Part of the issue about AMR in "popular medicine" is to discourage patients from seeking antibiotics for viral conditions that often present in very similar ways. Giving antibiotics for a virus is a pure social harm because they will do the patient no good and can only increase the risk of AMR.) This process thus matches, in a stylised way, the progress of many illnesses. Patients are contagious before they know they are ill. They become ill and seek treatment. Once they receive appropriate treatment, they will usually recover at least while antibiotics continue to work. (Of course, these stylised diseases are still very abstract. Parasites may need to be treated differently, as will sexually transmitted diseases and some illnesses require operations or other kinds of interventions than medication.)

The final complication, for a stage in the model that has not yet quite been reached, is that the disease is represented as a bit string of arbitrary "properties" (which determine which antibiotics work and *also* determine the symptoms). This is again plausible in a stylised way and allows relevant phenomena to be represented elegantly in the model. (For example, a virus and a bacterial infection can both present with fever and wheezing, so a laboratory test may be needed to avoid inappropriate antibiotic prescription). Thus the model is designed to respect the social dilemma at the core of the proposed research. Treatment can be based on symptoms alone or can use laboratory testing that identifies the actual bacteria involved. At the same time, the simulated medical professional faces the real challenge that the decision may be time-critical with the consequence that the patient might, albeit rarely, die.

The final element, which is the other "ology" in the title of the paper, is the way that the medical professional decides, based on the symptoms, what treatment to offer. This element is particularly simplified at present with assumptions of infallibility and free availability of antibiotics (that match any particular "bit string" of illness). This is the area of the model that will be most developed in later versions (but is also necessarily postponed by the research design of the project involving the need to achieve qualitative interviews with relevant medical professionals).

The ABM is designed to model a fixed geographical area (which could be thought of as Staphrica or Staphordshire) and, based on this view, is initialised by a relatively small number of agents (a vector) infected with the two illnesses commonly explored (one of which is a bacterium that needs antibiotic treatment and the other of which is a virus that will typically get better through natural resistance). The existence of disease "flare-ups" is easier to justify when thinking of bounded geographical areas as the process by which diseases continue to mutate and diffuse on a global scale would be much more challenging to model.

There is only really space to discuss one result in this paper, and this is somewhat in the nature of a baseline (though it still raises some interesting issues). Given the assumed severity, contagiousness and so on of the illnesses, how dangerous would they be if no medical intervention occurred at all? Conversely, even if medical treatment were "perfect" (in a way that can be defined within the model), would there still be fatalities? Fig. 9.1 shows the answer to this pair of "bracketing" questions.

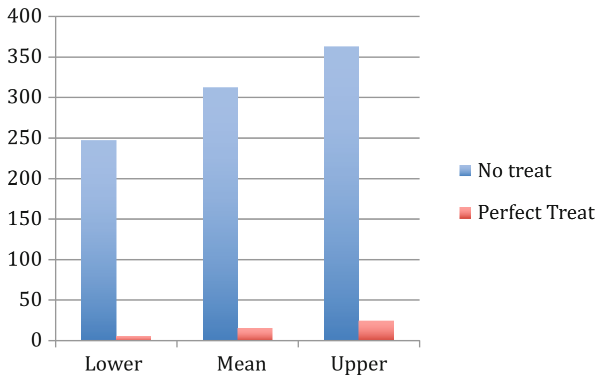


Fig. 9.1 Death tolls with no medical intervention and "perfect" intervention

These results come from a population of 500 agents, and it is therefore clear that, without medical intervention, the bacteria specified are extremely dangerous. There were ten runs in each simulation condition, and the three bars show the lowest and highest death tolls along with the mean for all ten runs to give a sense of stochastic variation. (In fact, it happens that nobody dies of the virus because it is strongly controlled by natural resistance.) Perfect treatment is defined as a situation where everyone seeks treatment as soon as they display symptoms, the medical professional can tell the virus from the bacteria with complete certainty and they treat immediately with exactly the needed antibiotic (which the patient takes without fail) for the particular bacteria (which involves a perfect bit string match). Importantly, even with perfect treatment, there is a death toll though it is clearly very much smaller than without treatment.

It is a commonplace that ABM may produce thought-provoking or counter-intuitive results and even this simple model appears to do so (though it should be noted that both terms have a strong subjective element and should thus be treated with caution). Firstly, it draws attention to the "voluntaristic" nature of medical help seeking in the system defined. Except in the event of communicable diseases that are very rare in the UK (it was interesting to see notices at surgeries during the Ebola outbreak telling potential patients to go away and phone in), treatment must be *sought*. This means that even if a patient seeks help as soon as they have symptoms and treatment is immediate and perfect, they may still be unlucky and die because their illness level happens to build up very fast stochastically even while the medicine is taking effect. In many countries, where treatment is expensive and difficult to access physically, illnesses may not present until they have become much more serious with a concomitant effect on death rates even assuming "optimal" treatment at the point of presentation.

The second point could almost be classed as rhetorical but is still important. There is a very strong medical culture of saving life at all costs (which creates issues with things like approved euthanasia). Apart from "driving" the social dilemma of AMR – patients here and now take precedence over the many patients "yet to come" (a general issue with broadly "utilitarian" reasoning) – this practice also leads to the view that the best outcome is "no death". This model suggests that this outcome is simply not attainable (and shows exactly why not in a particular kind of health system) but also shows that doing "something" is already hugely better than not intervening at all. Other preliminary experiments (not reported here) tend to show that while changes in prescribing behaviour do impact on lives saved to a small extent, most of the value added comes from deciding to intervene in the first place.

This leads to a final interesting point that concerns the importance of things *not* discussed in the model. Although, cross infection was deliberately ruled out in the model, obliging ill people to co-locate for treatment (either at a GP or in a hospital) is not an obviously advantageous strategy (both in terms of direct symptomatic infection and the possibility of transfer of bacteria with AMR), and it is interesting that the old model of "house calls" may have had medical advantages despite its apparent lack of "technical efficiency". In many countries, the cost and difficulty of seeking treatment has a profound effect on likely outcomes, but in fact, the same may be true (at least in specific circumstances) in the UK. Here the issue is not the great majority of cases successfully "fielded" by GP practices (which operate a vital and cost-effective "triaging" system) but the speed of response that can be achieved for the relatively few people who are much iller than the initial contact would suggest. Given even a stylised disease process, effective wait time (and the ability to negotiate it in the light of new information) may become a crucial element of health outcomes. This is true whether the waiting occurs at home, under treatment, in A&E, pending referral to a consultant or somewhere else. Even the stylised illness model suggests why this may be so.

9.4 Discussion and Conclusions

It has been remarked that peer review gives a wholly unrealistic sense of the degree of "finish" to be found in the research process. The main "result" of this research in progress has arguably been to "throw away" a significant part of the existing ABM dealing with microbiology! Nonetheless, it is clear that engaging with real research problems, even in the early stages, generates insights that unlikely to be developed by cogitation alone. In this section, I consider what this project has achieved and where it needs to go now.

As already suggested, the rapid development of a "chassis" ABM (which is nonetheless broadly "structurally valid"), based on an overview of existing literature and expert judgment, really does clarify thinking and provide a much better basis for focused discussion (particularly in collaborative research involving non-modellers). Without building the first model, it would not have been clear how incompatible it was with the idea of broad- and narrow-spectrum prescribing (despite its other many advantages) or exactly what questions to ask an expert to work out the relevant microbiological mechanisms for this. (As anecdotal support for an earlier claim, however, it is interesting how much experts disagree on the strengths and weaknesses of a model and how it would be most productive to develop it further. This situation, which I have observed in other projects, suggests that expert views must be treated with caution. Perhaps experts themselves need calibration and validation!)

Secondly, even the chassis model (as often claimed) throws up interesting and unexpected issues (about the role of "wait time" and the limited added value of particular interventions as opposed to intervention in general). If nothing else, this paper can be seen as another case study to support that broad claim about the virtues of ABM.

Thirdly, real problems create real challenges (which may not be the things that ABM is typically preoccupied with). The challenge of building models that "respect" (are empirically compatible with) the differing perspectives of microbiologist, patient and medical professional is not one that I have seen discussed in the literature and, interestingly, one that was discovered to be relevant to a rather different project at almost exactly the same time: Does it matter in effectively modelling people's attempts to avoid counterfeit drugs that the researchers, like the patients, cannot really get reliable information about the prevalence of such drugs? Thus, this paper encourages a focus on ABM challenges that arise from definite problems rather than the preoccupations of the method itself.

The final implication is one that is not fully realised but can already be sketched. It is not worthwhile to devise a methodology for ABM [10] unless it can actually be implemented. This project has been an attempt to put such a methodology to the test with associated benefits and pitfalls. Before data is collected (but while data collection is being designed), one or more "version 0" models should attempt to synthesise existing theories, establish face validity with experts and collaborators, achieve broad "structural validity" and get a sense of the feasibility of calibration and validation from existing data. Does what is already known even approximately permit the ABM strategy for establishing a "good" model? (It is now our regular experience that one needs to "fight for" validation data in ABM research and ensure that it exists, or can be collected, right from the outset of any project. The issue of rigorous research design [11, 12] generally seems to me to be neglected in ABM relative to other methods like statistics or qualitative interviewing. This would be a concrete example.) The data collection process is designed to simultaneously improve structural validity, calibration and validation using "gaps" identified in existing research as a framework. (This approach must necessarily be open-ended, and it should not be presumed that any single research project is capable of completing it, only contributing usefully to it.) In parallel, existing policy interventions and expert judgment must be used to ensure that the resulting ABM (once it has shown sufficient quality to be credible to policymakers) is "compatible with" the kinds of policies that are feasible and/or of interest. (Of course, there is no requirement that the model is used *only* to investigate such policies, but to be credible, it should *at least* be able to do that.) This project is thus part of a wider agenda to make a methodology of ABM workable in practice for policy-relevant problems.

References

1. Gao, D., Lietman, T.M., Porco, T.C.: Antibiotic resistance as collateral damage: the Tragedy of the Commons in a two-disease setting. Math. Biosci. **263**, 121–132 (2015). https://doi.org/10.1016/j.mbs.2015.02.007
2. Dyar, O.J., Huttner, B., Schouten, J., Pulcini, C.: What is antimicrobial stewardship? Clin. Microbiol. Infect. **23**, 793e798 (2017). https://doi.org/10.1016/j.cmi.2017.08.026
3. Chattoe-Brown, E., Colman, A.M., Jenkins, D.R., Krockow, E.M., Mehtar, S., Perera, N., Tarrant, C.: A Preliminary Description of an ABM to Investigate Collective Action Problems of Anti-Microbial Resistance (AMR) in Medical Settings. Draft paper, University of Leicester (2017)
4. Moreira, E., Costa, S., Aguiar, A.P., Câmara, G., Carneiro, T.: Dynamical coupling of multiscale land change models. Landsc. Ecol. **24**, 1183–1194 (2009). https://doi.org/10.1007/s10980-009-9397-x
5. Barnes, S., Golden, B., Wasil, E.: MRSA transmission reduction using agent-based modeling and simulation. INFORMS J. Comput. **22**, 635–646 (2010). https://doi.org/10.1287/ijoc.1100.0386
6. Koivumaa-Honkanen, H., Honkanen, R., Koskenvuo, M., Kaprio, J.: Self-reported happiness in life and suicide in ensuing 20 years. Soc. Psychiatry Psychiatr. Epidemiol. **38**, 244–248 (2003). https://doi.org/10.1007/s00127-003-0625-4
7. Etienne, M. (ed.): Companion Modelling: A Participatory Approach to Support Sustainable Development. Springer, Berlin (2014)
8. McPhee-Knowles, S.: Growing food safety from the bottom up: an agent-based model of food safety inspections. J. Artif. Soc. Soc. Simul. **18**, (2015). https://doi.org/10.18564/jasss.2717
9. Hotchkiss, J.R., Strike, D.G., Simonson, D.A., Broccard, A.F., Crooke, P.S.: An agent-based and spatially explicit model of pathogen dissemination in the intensive care unit. Crit. Care Med. **33**, 168–176 (2005). https://doi.org/10.1097/01.CCM.0000150658.05831.D2
10. Chattoe-Brown, E.: Agent based models. In Atkinson, P. A., Delamont, S., Hardy, M. A., Williams, M. (eds.) The Sage Encyclopedia of Research Methods. Sage (forthcoming 2019).
11. White, P.: Developing Research Questions: A Guide for Social Scientists. Palgrave Macmillan, Houndmills (2009)
12. Gorard, S.: Research Design: Creating Robust Approaches for the Social Sciences. Sage, Los Angeles (2013)

Chapter 10
Ethics-Based Cooperation in Multi-agent Systems

Nicolas Cointe, Grégory Bonnet, and Olivier Boissier

Abstract In the recent literature in Artificial Intelligence, ethical issues are increasingly discussed. Many proposals of ethical agents are made. However, those approaches consider mainly an agent-centered perspective, letting aside the collective dimension of multi-agent systems. For instance, when considering cooperation among such agents, ethics could be a key issue to drive the interactions among the agents. This paper presents a model for ethics-based cooperation. Each agent uses an ethical judgment process to compute images of the other agents' ethical behavior. Based on a rationalist and explicit approach, the judgment process distinguishes a theory of good, namely, how values and moral rules are defined, and a theory of right, namely, how a behavior is judged with respect to ethical principles. From these images of the other agents' ethics, the judging agent computes trust used to cooperate with the judged agents. We illustrate these functionalities in an asset management scenario with a proof of concept implemented in the JaCaMo multi-agent platform.

Keywords Computational ethics · Ethical judgment · Agent cooperation

N. Cointe (✉)
Department of Engineering Systems and Services, Delft University of Technology, Delft, Netherlands
e-mail: nicolas.cointe@tudelft.nl; nicolas.cointe@emse.fr

G. Bonnet (✉)
GREYC UMR CNRS 6072, UNICAEN, ENSICAEN, Normandy University, Caen, France
e-mail: gregory.bonnet@unicaen.fr

O. Boissier (✉)
MINES Saint-Étienne, CNRS, Laboratoire Hubert Curien, UMR 5516, Université de Lyon, Lyon, France
e-mail: Olivier.Boissier@emse.fr

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_10

10.1 Introduction

The increasing use of autonomous agents in real-world application raises many issues. Besides achieving goals optimally, the ethical and moral dimensions of an agent's decisions should be considered in their reasoning. For instance, in a multi-agent-based asset management, many models are available to evaluate the potential profit of an investment. Nevertheless it is still not possible for an agent to judge the morality and ethics of this investment. Moreover, the heterogeneity of these elements raises many issues when agents need to collaborate with other agents while respecting their own ethics. Given its own ethics, how could an agent compute its ethical conformity with other agents from their observed behaviors? How could this agent decide to trust agents based on these images? To answer such questions, we extend our previous work on ethical judgment process [15]. This work proposed an architecture enabling agents to judge either their own behavior or the other agents' one. However, it does not allow to judge a behavior on a time interval. Consequently, we extend this architecture with a mechanism to build an image of the other agents by computing and aggregating evaluations of these judgments. Then, we propose a mechanism allowing an agent to decide on trusting the other agents for their morality or their ethics. Finally, the components of this proposal are instantiated to the asset management domain to demonstrate their use. Hence, Sect. 10.2 introduces and describes the concepts of trust and ethics we use. Then, Sect. 10.3 shows how ethical judgment can be used to depict in terms of ethics the others' behavior. Sect. 10.4 presents the construction and use of trust, and finally we illustrate their use in Sect. 10.5.

10.2 Foundations

We introduce in this section the necessary concepts to deal with ethics-based cooperation in multi-agent systems. Section 10.2.1 presents the concept of *trust* as a way to ground interaction and cooperation among agents. Section 10.2.2 introduces ethics and shows how it is related to trust. Finally, Sect. 10.2.3 synthesizes the requirements for defining an ethical-based cooperation in a multi-agent systems.

10.2.1 Trust in Multi-agent Systems

In decentralized and open systems, a way to deal with unreliable or unknown agents is to use trust [12, 17, 32]. Trust allows the agents to assess the interactions they observe or they make in order to decide if interacting with a given agent is a priori acceptable. This acceptance notion means that the investigated agent behaves well and is reliable according to the investigator criteria. Many definitions of trust exist, but, in accordance with [12], we consider *trust* as *a disposition to cooperate with* *a trustee*. It can be used as a condition to perform actions as delegating actions, sharing resources and information, or any kind of cooperation. To build trust, the agents first build an image of the investigated agents [17]. An *image* is *an evaluative belief that tells whether the target is good or bad with respect to a given behavior*. In the literature, images are aggregated from the experiences, i.e., the observed behavior of the target agent and its consequences. We can distinguish two kinds of approaches:

- Statistical images [1, 10, 18, 26, 36] where the image is a quantitative aggregation of feedbacks about interactions. This aggregation estimates the trends of an agent to behave well from another agent's point of view. It can be represented by Bayesian networks, Beta density functions, fuzzy sets, Dempster-Shafer functions, and other quantitative formalisms.
- Logical images [11, 12, 28, 35] where the image is a mental state rooted in every cooperation action that is produced by interactions. A persistent image allows to infer trust beliefs that can be used as preconditions to cooperate.

An agent can lack of observations in order to build a correct image of a target. A way to deal with this issue is to use reputation [24, 31]. It consists in using third-party agents' image of the target (that can depend on the initial agent's image has about the third parties) in order to assess a collective point of view about the target. Both images and reputations are used to lead to a trust action [32]. Most of the time, trust is dynamic and changes with respect to the evolution of images and reputations.

10.2.2 Ethical Behaviors

Due to the lack of formal definitions of the components of morals and ethics in the literature, we based our definition of moral theories on two components [3, 29, 34]: theory of the good (or morals) and theory of the right (or ethics). The theory of the good describes what actions are morally good, while the theory of the right describes what an agent ought to do according to the theory of the good. It means that an ethical behavior is grounded by theory of the right to conciliate the theory of the good and the desires and the capacities of the agent. More precisely:

- A *theory of the good* is a set of moral rules and values which allow to assess the goodness or badness of an action itself. Moral rules give moral valuations to behaviors (e.g., "Lying is evil" or "Being honest is good"), and values give them more abstract qualification (e.g., "Telling what we believe is being honest").
- A *theory of the right* uses a set of ethical principles to recognize a fair or, at least, acceptable option in comparison with the other available actions in a given situation. Philosophers proposed various ethical principles, such as Kant's Categorical Imperative [23] or Thomas Aquinas' Doctrine of Double Effect [27]. For example, even if stealing can be considered as immoral (regarding Divine Commands), some philosophers agree that it is acceptable for starving people to rob food (regarding Doctrine of Double Effect).

Thus, theories of the right follow from and rely on theories of the good. Even if this distinction is still debatable due to the diversity of existing contradictory theories in philosophy and psychology, it offers a sound framework to define moral and ethics. Interestingly, being moral (namely, acting according some moral rules or values) or ethical (acting according a given ethical principle) is a behavior characterization such as being reliable is in trust systems. Consequently, it can be interesting to define a trust notion based on moral or ethical behaviors, which can enhance cooperation.

10.2.3 Requirement for Ethical-Based Cooperation

Works dealing with ethical behaviors in autonomous agents often focus on modeling moral reasoning [7, 20, 21, 33] as a direct translation of some well-known moral theories or on modeling moral agency in a general way [6, 25]. However, those works do not clearly make the distinction between theory of the good and theory of the right. Some other works deal with ethical agent architecture. In the literature, we find *implicit ethical architectures* [4, 5] which design the agent's behavior either by implementing for each situation a way to avoid potential unethical behaviors or by learning from human expertise. We also find *cognitive ethical architectures* [13–16] which consist in full explicit representations of each component of the agent, from the classical beliefs (information on the environment and other agents), desires (goals of the agent), and intentions (the chosen actions) to some concepts as heuristics or emotional machinery. However, all those approaches do not take into account the collective dimension of agent systems, apart [30] which considers morals as part of agent societies. More precisely, the architecture given in [15] makes a clear separation between theory of the good and theory of the right and provides beliefs on various components of moral theories (moral rules, values, or ethical principles, for instance). Moreover, the architecture given in [30] allows – but without operationalization – moral facts (judgments over other agents or blames, for instance) to be viewed as beliefs that can be used in the agents' decisions. In order to build ethics-based cooperation, we need an operational model of ethical judgment such as proposed in [15]. Inspired by [30], we reuse and extend this model with beliefs on moral and ethical images of other agents. We use those image beliefs to build trust beliefs to drive a cooperation based on morals or ethics.

10.3 Judgments Building

This section explains how the judgment process introduced in [15] is used (see Sect. 10.3.1) and presents how an agent computes its own qualitative representation of the ethics (see Sect. 10.3.2) and morals (see Sect. 10.3.3) of the other agents, regarding the judging agent's knowledge on morals and ethics.

10.3.1 Judging Other Agents

Let us consider the judgment process introduced in [15]. It is used to generate a set of rightful actions $\mathcal{A}_r$ for a given situation, regarding a set of knowledge. The judgment process depicted in Fig. 10.1 is organized into three parts: (*i*) awareness and evaluation process, (*ii*) goodness process, and (*iii*) rightness process. In this architecture, *knowledge base* is a priori knowledge given to the agents, and *mental states* are knowledge inferred from their knowledge bases and other mental states thanks to *functions*. Since this judgment process may use sets of knowledge issued from another agent, all these sets are indexed with an agent $\mathrm{a} \in \mathbb{A}$ (e.g., $\mathcal{A}_{r_\mathrm{a}}$) with $\mathbb{A}$ the set of the agents. When a is the agent executing the process, it judges its own behavior. When a is different, the agent judge a's behavior.

Awareness and evaluation processes The awareness process produces $\mathcal{D}_\mathrm{a}$ the set of desires and $\mathcal{B}_\mathrm{a}$ the set of beliefs of a with a situation assessment SA of the current world W. The *evaluation* process evaluates the set of actions $\mathcal{A}_\mathrm{a}$ (actions are pairs of conditions and consequences bearing on desires and beliefs) in terms of desirable ($\mathcal{A}_{d_\mathrm{a}}$) and executable ($\mathcal{A}_{c_\mathrm{a}}$) actions with respect to $\mathcal{D}_\mathrm{a}$ and $\mathcal{B}_\mathrm{a}$. Those evaluations are done, respectively, by DE and CE functions. In the sequel, we call *contextual knowledge* of a (CK_a), the union of $\mathcal{B}_\mathrm{a}$ and $\mathcal{D}_\mathrm{a}$.

Goodness process The *goodness process* use the ME function to compute moral actions $\mathcal{A}_{m_\mathrm{a}}$ given a's contextual knowledge CK_a, actions A_a, value supports VS_a, and moral rules MR_a. Moral actions are actions that, in the situations of CK_a, promote or demote the moral values of VS_a. A *value support* is a tuple $\langle s, v\rangle \in VS_\mathrm{a}$ where $v \in \mathcal{O}_v$ is a moral value and $s = \langle \alpha, w\rangle$ is the support of this moral value where $\alpha \in A_\mathrm{a}$, $w \subset \mathcal{B}_{ai} \cup \mathcal{D}_\mathrm{a}$. $\mathcal{O}_v$ is the set of moral values used in the system. A *moral rule* is a tuple $\langle w, o, m\rangle \in MR_\mathrm{a}$ meaning in the situation w acting as o is morally evaluated as m. The situation $w \in 2^{CK_\mathrm{a}}$ is a conjunction of beliefs and desires. The act o is either an action $\alpha \in A_\mathrm{a}$ or a moral value $v \in \mathcal{O}_v$ (meaning acting as α or in a way to promote v). Finally, $m \in \mathcal{O}_m$ is the moral valuation. For instance, $\mathcal{O}_m = \{$`moral`, `amoral`, `immoral`$\}$ with a *total order* (e.g., `moral` is a higher moral valuation than `amoral`, which is higher than `immoral`). In the

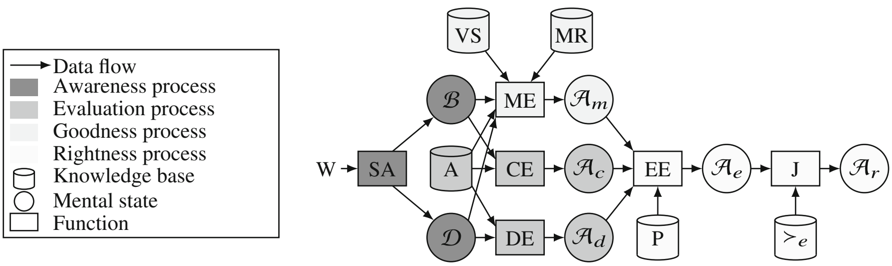


Fig. 10.1 Ethical judgment process as depicted in [15]

sequel, moral rules MR_{a}, value support VS_{a}, and values O_v, knowledge used in the goodness process of the agent a, are referred as the *goodness knowledge* (GK_{a}).

Rightness process Finally, the rightness process assesses the rightful action $\mathcal{A}_{r_{\text{a}}}$ from the sets of possible $\mathcal{A}_{c_{\text{a}}}$, desirable $\mathcal{A}_{d_{\text{a}}}$, and moral $\mathcal{A}_{m_{\text{a}}}$ actions based on *ethical principles* P_{a} to conciliate these sets of actions according to ethical preference relationship $\succ_{e_{\text{a}}} \subseteq P_{\text{a}} \times P_{\text{a}}$. An *ethical principle* $p \in P_{\text{a}}$ is a function which evaluates if it is *right* or *wrong* to execute a given action in a given situation regarding a philosophical theory. It describes the rightness of an action with respect to its belonging to $\mathcal{A}_{c_{\text{a}}}$, $\mathcal{A}_{d_{\text{a}}}$, and $\mathcal{A}_{m_{\text{a}}}$ in a given situation of CK_{a}. It is defined as $p : 2^{\mathcal{A}_{\text{a}}} \times 2^{\mathcal{B}_{\text{a}}} \times 2^{\mathcal{D}_{\text{a}}} \times 2^{MR_{\text{a}}} \times 2^{V_{\text{a}}} \rightarrow \{\top, \bot\}$. Given a set of actions issued of the ethic evaluation function EE that applies the ethical principles, the judgment function J is the last step which selects the set of rightful actions to perform, considering the set of ethical preferences $\succ_{e_{\text{a}}}$ defining a total order on the ethical principles. In this judgment function, the rightful actions are the ones that satisfy the most preferred principles in a lexicographic order. In the sequel, ethical principles P_{a} and preferences $\succ_{e_{\text{a}}}$ are referred as the rightness knowledge (RK_{a}).

10.3.2 Judging Ethical Conformity of Behaviors

We extend now the previous judgment process to judge the ethics and morality of the behavior between t_0 and t of an agent a′. Inspired from [30] which considers beliefs on moral facts, the judgment process produces now beliefs (`ethical_conformity`, `moral_conformity`) stating the conformity to ethical principles or moral rules and values, which can be used in the agent's reasoning. Before defining these beliefs, let us define first an agent's behavior as follows:

Definition 1 (Behavior) The *behavior* $b_{\text{a}',[t_0,t]}$ of an agent a′ on the time interval $[t_0, t]$ is the set of actions α_k that a′ executed between t_0 and t as $0 \leqslant t_0 \leqslant t$.

$$b_{\text{a}',[t_0,t]} = \{\alpha_k \in A : \exists t' \in [t_0, t] \text{s.t.} \texttt{done}(\text{a}', \alpha_k, t')\}$$

where $A = \bigcup_{\text{a}_i=\text{a}_1}^{\text{a}_n} \mathcal{A}_{\text{a}_i}$ is the set of all available actions for all agents and $\texttt{done}(\text{a}', \alpha_k, t')$ means that α_k has been executed[1] by a′ at time t'.

An agent a can judge the conformity of an action α_k executed by another agent a′ with respect to its own goodness and rightness knowledge.

Definition 2 (Ethical conformity) An action α_k is said to be *ethically conform* with respect to the judging agent a's contextual knowledge (CK_{a}), goodness knowledge GK_{a}, and rightness knowledge RK_{a} at time t', noted $\texttt{ethical\_conformity}(\alpha_k, t')$, if and only if α_k is in the set of rightful actions

[1] A behavior can deal with concurrency: several actions can have been done at the same time.

$\alpha_k \in \mathcal{A}_{r_a}$ computed by the ethical judgment J_a of the judging agent a, based on $[CK_a, GK_a, RK_a]$ at time t'.

Let us notice that the ethical conformity of an action can be applied to judging agent's actions or to actions executed by another agent and observed by the judging agent. This ethical conformity can be judged with respect to the judging agent's contextual, goodness, and rightness knowledge. It can be judged also with respect to the rightness or goodness knowledge of another agent as long as the judging agent has a representation of these knowledge. Then, the ethical conformity is used to compute the set EC^+ of ethically conform (resp. the set EC^- of non-ethically conform) actions of the judged agent a′'s behavior $b_{a',[t_0,t]}$ between t_0 and t:

$$EC^+_{b_{a'},[t_0,t]} = \{\alpha_k \in b_{a',[t_0,t]} \wedge t' \in [t_0, t] \text{s.t.} \mathtt{done}(a', \alpha_k, t') \wedge \mathtt{ethical\_conformity}(\alpha_k, t')\}$$
$$EC^-_{b_{a'},[t_0,t]} = \{\alpha_k \in b_{a',[t_0,t]} \wedge t' \in [t_0, t] \text{s.t.} \mathtt{done}(a', \alpha_k, t') \wedge \neg \mathtt{ethical\_conformity}(\alpha_k, t')\}$$

These two sets provide information on the behavior of the judged agent and its compliance with the ethics of the judging agent. Nevertheless, it cannot assess why an observed behavior is judged as unethical. Indeed, the reason can be a difference between the judging and the judged agents' theory of the right, theory of the good, or the assessment of the situation. In the sequel, we will denote: $EC_{b_{a'},[t_0,t]} = EC^+_{b_{a'},[t_0,t]} \cup EC^-_{b_{a'},[t_0,t]}$

10.3.3 Judging Moral Conformity of Behaviors

The moral conformity of an action with respect to a given moral rule is realized regarding a moral threshold $mt \in MV$ and a situation assessment.

Definition 3 (Moral conformity) An action α_k is said to be *morally conform* at time t' with respect to the judging agent a's contextual knowledge CK_a and goodness knowledge GK_a, considering the moral rule $mr \in MR_a$, moral threshold $mt \in MV_a$, and noted $\mathtt{moral\_conformity}(\alpha_k, mr, mt, t')$, if and only if α_k belongs to $\mathcal{A}_{m_a}$ with a moral valuation greater or equal to mt, given the considered moral rule mr, CK_a, and GK_a at time t'.

Similar to the ethical conformity, the moral conformity of an action defines the set MC^+ (resp. MC^-) of morally conform (resp. non morally conform) actions of a′'s behavior $b_{a',[t_0,t]}$ of during $[t_0, t]$ with respect to mr and mt:

$$MC^+_{b_{a',[t_0,t]},mr,mt} = \{a_k \in b_{a',[t_0,t]} \wedge t' \in [t_0, t] \text{s.t.} \mathtt{done}(a', a_k, t') \wedge \mathtt{moral conformity}(a_k, mr, mt, t')\}$$
$$MC^-_{b_{a',[t_0,t]},mr,mt} = \{a_k \in b_{a',[t_0,t]} \wedge t' \in [t_0, t] \text{s.t.} \mathtt{done}(a', a_k, t') \wedge \neg \mathtt{moral conformity}(a_k, mr, mt, t')\}$$

As the *ethical evaluation* function returns a boolean, *moral evaluation* associates to the action a moral valuation that should be compared with a threshold to be considered conform or not. We can generalize the above evaluation of the moral conformity with respect to a moral rule to a set of moral rules, considering the possibility to define a subset ms of moral rules $ms \subseteq MR_{\text{a}}$. Such a set ms represents a cluster of rules such as rules based on some moral values, rules concerned by particular situations, and so on. In the sequel, we denote: $MC_{b_{\text{a}'},ms,mt,[t_0,t]} = MC^{+}_{b_{\text{a}'},ms,mt,[t_0,t]} \cup MC^{-}_{b_{\text{a}'},ms,mt,[t_0,t]}$.

10.4 Trust Within Ethical Behavior

In this section, conformity beliefs defined previously are used to compute the images of other agents (see Sect. 10.4.1). We introduce first how these images are used to build trust in Sect. 10.4.2. Then, Sect. 10.4.3 provides hints about how to use it.

10.4.1 Ethical and Moral Images of an Agent

Following Sect. 10.2.1, the *ethical* and *moral* images of an agent are evaluative beliefs that tell whether another agent has a conform behavior or not with respect to a given rightness (RK) and goodness (GK) knowledge.

Definition 4 (Ethical Image (resp. Moral Image)) An *ethical image* (resp. *moral image*) of an agent a′[2] is the judgment of the behavior $b_{\text{a}',[t_0,t]}$ of that agent in a situation with respect to an ethics (resp. to set of moral rules ms and a moral threshold mt), regarding the contextual CK, goodness GK and rightness RK knowledge of another agent a. This image states a conformity valuation $cv \in CV$, where CV is an ordered set of *conformity valuation*.[3] They are noted as `ethical_image`(a′, a, cv, t_0, t) and `morality_image`(a′, a, cv, ms, mt, t_0, t)

Indeed, while an agent can only have a single ethical image of other agents, it can have several moral images of the same agents depending on the chosen ms and mt. To build these images, an agent a uses two aggregation functions EA and MA applied, respectively, on evaluated actions regarding ethics $EC_{b_{\text{a}'},[t_0,t]}$ and regarding moral $MC_{b_{\text{a}'},[t_0,t]}$. Both aggregation functions compute the ratio of the weighted sum of positive evaluations with respect to ethics and with respect to morals. The weight of each action corresponds to a criterion (e.g., the time past from the date of the evaluation, the consequences of the action, and so on).

Definition 5 (Ethical aggregation function) The function $EA : 2^{\mathcal{A}} \rightarrow [0, 1]$ is defined such that $EA(EC_{b_{\text{a}'},[t_0,t]}) =$

[2] Let's notice that in the definition of these images, the second parameter refers to an agent. It means that the image is built with respect to the knowledge of this agent. The first parameter refers to the considered agent's behavior.

[3] As for morals, conformity valuations are for instance { improper, neutral, congruent }.

$$\sum_{\alpha_k \in EC^+_{b_{a'},[t_0,t]}} weight(\alpha_k) \;/ \sum_{\alpha_k \in EC_{b_{a'},[t_0,t]}} weight(\alpha_k)$$

Definition 6 (Moral aggregation function) The function $MA : 2^{\mathcal{A}} \rightarrow [0, 1]$ is defined such that $MA(MC_{b_{a'},[t_0,t]}) =$

$$\sum_{\alpha_k \in MC^+_{b_{a'},[t_0,t]}} weight(\alpha_k) \;/ \sum_{\alpha_k \in MC_{b_{a'},[t_0,t]}} weight(\alpha_k)$$

In order to transform the quantitative evaluation into a qualitative one, every conformity valuation is associated to an interval in the range of the ethical and moral aggregation functions. Once the conformity valuation computed, the associated beliefs `moral_image`$(a', a, ms, mt, cv, t_0, t)$ or `ethical_image`(a', a, cv, t_0, t) are produced. For instance, if `congruent` conformity evaluation is defined in [0.75, 1], the behavior of an agent is considered as ethical if $EA(EC_{b_{a'},[t_0,t]}) \geq 0.75$. Finally, those images can be used to influence interactions by building trust relationships, or to describe the morality of interactions, depending on the behavior of the others.

10.4.2 Building Trust Beliefs

According to the information on the moral and ethical images, an agent can decide to trust others or not. Trust can be absolute (trust in the rightness of the others' behavior) or relative to a set of moral rules (trust in their responsibility, carefulness, obedience to some sets of rules, and so on). We define two internal epistemic actions, with respect to ethical and moral images, respectively, that build beliefs on trust.

Definition 7 (Trust function) The *ethical trust function* TB^e_a (resp. *moral trust function* TB^m_a) is $TB^e_a : \mathbb{A} \rightarrow \{\top, \bot\}$ (resp. $TB^m_a : \mathbb{A} \times 2^{\mathcal{MR}_a} \times MV_a \rightarrow \{\top, \bot\}$).

Here, those trust functions are abstract and must be instantiated. For example, when an agent a computes that the behavior of another agent a′ is conform with CK_a, GK_a and RK_a (i.e., the ethical image), the ethical trust function produces a belief `ethical_trust`(a', a). Similarly, when the agent a computes that a′'s behavior is conform with ms (i.e., the moral image of its behavior regarding ms is at least mt), the moral trust function produces a belief `moral_trust`(a', a, ms, mt).

10.4.3 Ethical Trusting

Beliefs on images and trust can be used as a part of the context to evaluate the morality and ethics of an action. To this end, we can express that the morality of an action that affect other agents depends on their image. Firstly, ethical and moral trust can enrich the description of the moral rules or values. It is useful to represent that the others' behavior can have an impact on how a context is qualified. For instance, the *responsibility* value may be supported by delegating actions to ethically trusted agents only. Here, responsibility is defined as the capability to act safely with the appropriate agents. We can also explicitly express it is not responsible to delegate something to an agent known for its unethical behavior. Secondly, specific moral trust beliefs can be used as elements of moral rules. For instance, assuming a *honesty* moral value and its value supports, an agent can express the moral rule: "It is immoral to not behave honestly towards an agent who is trusted as being honest." Here, "who is trusted as being honest" can be modeled by a `moral_trust` belief where the associated moral rules *ms* are all rules that refer to honesty. Finally, as evaluating and judging others are actions, it is also possible to evaluate their morality or ethics. For instance, *tolerance* as a moral value might be supported by building an image on the others with a low moral threshold until the sets $EC_{\mathrm{a}',[t_0,t]}$ or $MC_{\mathrm{a}',[t_0,t]}$ are significant enough. The choice of the thresholds, the weights, and the conversion of the aggregation into a conformity valuation can also be a way to represent various types of trust. As another example, *forgiveness* can a value supporting high weights on the most recent observations. It can allow then to specify an ethics of trust as "It is immoral to build trust without tolerance and forgiveness" [22].

10.5 Proof of Concept

This section illustrates how the elements presented in the previous sections have been implemented in a multi-agent system. We use the JaCaMo platform [8] where the agents are programmed in BDI architecture using the Jason language and the shared environment is programmed with workspaces and artifacts from the CArtAgo platform. The complete source code is available on our website.[4] The environment is a simulated asset market where assets are quoted, bought, and sold by autonomous agents. Section 10.5.1 introduces the features of our application. Morals and ethics are defined in Sect. 10.5.2. Images and trust building are shown in Sect. 10.5.3.

10.5.1 Asset Market Modeling

Trading assets lead to several practical and ethical issues.[5] This is all the more important in automated trading as decisions, made by autonomous agents to whom human users delegate the power to sell and buy assets, have consequences in real life [19]. As shown by [9], some investment funds are interested to make socially

[4] http://www.nicolascointe.eu/projects/ethical_market_simulator

[5] http://sevenpillarsinstitute.org/

responsible and ethical trading, and they are growing and taking a significant position on the market. However, whereas the performance of such funds can be measured objectively, their ethical quality is more difficult to assess as it depends on the values of the observer. In this proof of concept, we consider a market where autonomous trading agents can manage portfolios in order to sell or buy assets. Assets types are currencies – i.e., money – and equity securities, i.e., part of a company's capital stock. A market is represented as a tuple ⟨ `name`, `id`, `type`, `matching` ⟩ with the name of the market `name`, a unique identifier `id`, the type of exchanged assets `type`, and the algorithm used to store and execute orders `matching`. On the market, each agent can execute `buy`, `sell`, or `cancel` orders. They, respectively, correspond in exchanging an equity for a currency, exchanging a currency for an equity, and canceling an exchange order that has not been executed yet. Each equity is quoted in a state-of-the-art central limit order book (CLOB) [2] algorithm. By observing the market, the agents get beliefs on the market. Agents perceive each minute the volume (the quantity of exchanged assets), two moving means, representing the average price on the last 20 minutes and on the last 40 minutes, the standard deviations of prices on the last 20 minutes, the closing prices on this period, and the up and down Bollinger Bands (the average prices ± twice the standard deviations). Agents have also beliefs on the orders added and stored in the CLOB and their execution. The general form of all those beliefs is, respectively:

```
indicators(Date,Mktplace,Asset,Close,Volume,Intensity,Mm,Dblmm,
                                                    BUp, BDown)
onMarket(Date,Agent,Portfolio,Marketplace,Side,Asset,Volume,
                                                         Price)
executed(Date,Agent,Portfolio,Marketplace,Side,Asset,Volume,
                                                         Price)
```

A set of beliefs `own(PortfolioName,Broker,Asset,Quantity)` updated in real time represent the agents' portfolio. By reasoning on those beliefs as a contextual knowledge CK, an agent is able to infer the feasibility of passing a buy or sell order (simply by verifying if its own portfolio contains the assets to exchange) to produce $\mathcal{A}_p$. It can also reason on the desirability of these actions to produce $\mathcal{A}_d$. To this end, we implemented a simple but classical method of trading decision-making based on comparisons between the Bollinger Bands and the moving means. Introduced in our experiment are two types of agents: (1) *zero-intelligence agents* make random orders (in terms of price and volume) on the market to generate activity and simulate the "noise" of real markets. Each of them is assigned to one or every assets; (2) *ethical agents* implement the ethical judgment on their own actions as a decision process to make their decisions. They have a simple desirability evaluation function to speculate: if the price of the market is going up (the shortest moving mean is over the other one), they buy the asset; otherwise, they sell it. If the price goes out of the Bollinger Bands, these rules are inverted.

10.5.2 Ethical Settings

We consider that the ethical agents are initialized with a particular set of beliefs about activities of the companies (e.g., an energy producer using nuclear power plants) and some labels about their conformity with international standards (e.g., an electric infrastructure producer labeled FSC). Those beliefs are important to assess how it is moral to trade a given asset based on the company's activities. Indeed, to provide information on the morality of acting on a financial market, we implemented moral values and moral rules directly inspired from the literature available online.[6] The ethical agents know a set of organized values: for instance, "environmental reporting" is considered as a subvalue of "environment." Values are represented as:

```
value("environment").
subvalue("promote_renewable_energy","environment").
subvalue("envirnmt_reporting","environment").
```

Agents have a set of value supports as "trading assets of nuclear energy producer is not conform with the subvalue *promotion of renewable energy*", represented as:

```
valueSupport(buy(Asset,_,_,_),"envirnmt_reporting"):
                                        -label(Asset,"FSC").
```

Agents are also equipped with moral rules stating the morality of environmental considerations. For instance, "It is moral to act in conformity with the value *environment*" is simply represented as:

```
moral_eval(X,V1,moral):- valueSupport(X,V1) & subvalue
                                          (V1,"environment").
moral_eval(X,"environment",moral):- valueSupport(X,
                                              "environment").
moralSet("environment","value_environment").
```

We declare in the last line this moral rule as an element of a set of moral rules related to environmental values (in order to build images). Here, an ethical agent is able to infer, for instance, that, regarding its beliefs and this goodness knowledge, trading the asset of the FSC labeled company is moral while trading the asset of the nuclear energy producer is both moral and immoral. Thus, the agent needs a rightness knowledge to discriminate if it is right or wrong to trade the second assets. Finally, ethical agents are equipped with ethical principles, such as the Aristotelian ethics (inspired from [21]) and more simple principles such as considering `perfectAct` "It is rightful to do a possible, moral and desirable action," the nonshaming desire `desireNR` "It is rightful to do a possible, not immoral and desirable action," and the moral duty `dutyNR` "It is rightful to do a possible, moral and not undesirable action." Please see directly the file `rightness_process.asl` for more details.

[6] http://www.ethicalconsumer.org/

10.5.3 Image and Trust Building

Each time an action is executed on the market (i.e., a buy order matches with a sell order), the agents receive a message and evaluate their image of the agents implied in the transaction. As said in the previous section, evaluating the conformity of behaviors and building the image and the trust beliefs are actions. Thus, they are implemented as Jason plans. In the sequel, we will detail moral trust building. Ethical trust building is based on the same ideas. The following plan evaluates the conformity of the action with each moral rule of the set `MSet` and increments the value `X` stored in the belief `moralAggr(Agent,MSet,X)`. In this implementation, we use a linear aggregation, (i.e., it associates the same weight with each action). Then, a conformity valuation is computed regarding the proportion of conform actions in order to build the image. We use here three conformity valuations (arbitrary `neutral` for an aggregated ratio in $[0.4, 0.6[$, `improper` if lower and `congruent` if higher). Finally, when the conformity valuation crosses a trust threshold, a plan updates the trust belief in the judged agent regarding the set of moral rules.

```
+!trust : moralImageOf(Agent,MoralSet,ConformityValuation)
  & trustThreshold(Threshold) & not trust(Agent,MoralSet)
  & not tOrderOnConformityValuation(Threshold,Conformity
                                                     Valuation)
  <- +trust(Agent,MoralSet); !trust.
```

Similarly, we have implemented a plan for ethical conformity which stores the number of conform and nonconform actions regarding the rightness knowledge, a plan for ethical image building, and a plan for ethical trust building.

10.5.4 Results

Figure 10.2 shows the evolution of the ethical aggregations computed by an ethical agent on the others' behaviors. In this simulation, three groups of ethical agents are created with three different theories of good. Let us notice that the judge agent evaluates the behavior of both ethical and zerointelligence agents. Our model only evaluates the conformity of an observed behavior with an ethics, without trying to understand or reason on the intentions of the other agents. As expected, the ethical agents obeying the same ethics stay at a similar value (thick lines). The agents obliged to generate activity on such assets stay at 0.0 or 1.0 because they respectively can't do moral or evil actions regarding the judge's point of view. All the other agents slowly converge toward a value depending on their behavior.

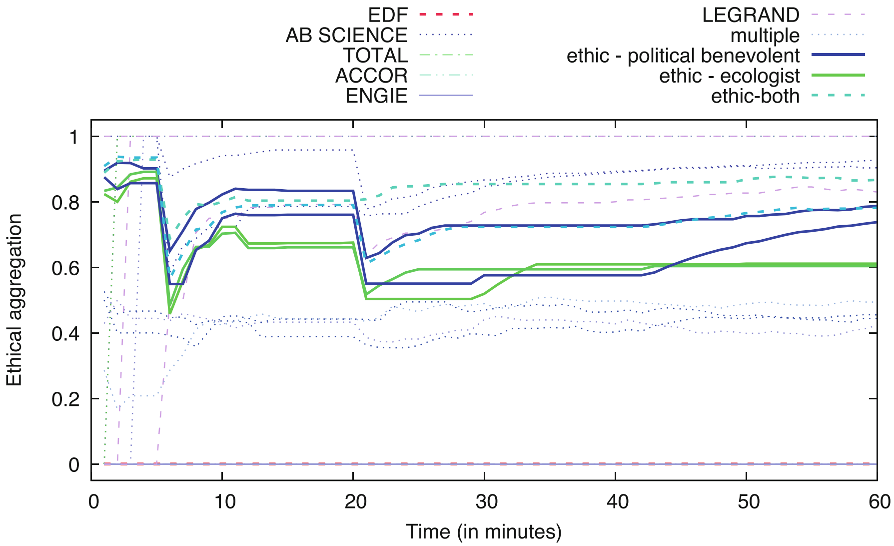


Fig. 10.2 Evolution of the output of an ethical aggregations functions

10.6 Conclusion

In this paper, we mapped a model of ethical judgment process into a BDI agent model and defined mechanisms to build images depicting the conformity of a behavior with respect to an ethics or morals. We demonstrated how agents can use these images to decide about trusting other agents in order to cooperate and delegate actions. A proof of concept shows how this model has been implemented in a BDI framework to be used in an asset management application. From a modeling point of view, this proposal addresses the problem of measuring how far from an ethics or a moral theory a behavior is, especially when ethics and morals lie in the hidden personal motivations and rules of a set of heterogeneous agents. With this model, agents may know which moral rules or values as well as ethics are concerned by this proximity. Thanks to the expressiveness of our model for ethics and morals, we envision as a future work to represent a large set of moral values that take the images of the others into account in their descriptions, such as forgiveness or intransigence.

References

1. Abdul-Rahman, A., Hailes, S.: Supporting trust in virtual communities. In: 33th IEEE International Conference on Systems Sciences, pp. 1–9 (2000)
2. Aldridge, I.: High-frequency trading: a practical guide to algorithmic strategies and trading systems, vol. 459. Wiley, Somerset (2009)

3. Alexander, L., Moore, M.: Deontological ethics. In: Zalta, E.N. (ed.), The Stanford Encyclopedia of Philosophy, Spring 2015 edition, Metaphysics Research Lab, Stanford University (2015)
4. Anderson, M., Anderson, S.L.: Toward ensuring ethical behavior from autonomous systems: a case-supported principle-based paradigm. Ind. Robot. **42**(4), 324–331 (2014)
5. Arkin, R.C.: Governing lethal behavior: Embedding ethics in a hybrid deliberative/reactive robot architecture. In Proceedings of the 3rd ACM/IEEE international conference on Human robot interaction, 121–128 (2008, March)
6. Arkoudas, K., Bringsjord, S., Bello, P.: Toward ethical robots via mechanized deontic logic. In: AAAI Fall Symposium on Machine Ethics, pp. 17–23 (2005)
7. Berreby, F., Bourgne, G., Ganascia, J.-G.: Modelling moral reasoning and ethical responsibility with logic programming. In: LPAR Proceedings, pp. 532–548 (2015)
8. Boissier, O., Bordini, R.H., Hübner, J.F., Ricci, A., Santi, A.: Multi-agent oriented programming with JaCaMo. Sci. Comput. Program. **78**(6), 747–761 (2013)
9. Bono, S., Bresin, G., Pezzolato, F., Ramelli, S., Benseddik, F.: Green, social and ethical funds in Europe. Technical report, Vigeo (2013)
10. Carbo, J., Molina, J., Davila, J.: Comparing predictions of SPORAS vs. a fuzzy reputation agent system. In: Conference on Fuzzy Sets and Fuzzy Systems, pp. 147–153 (2002)
11. Carter, J., Bitting, E., Ghorbani, A.: Reputation formalization for an information-sharing multi-agent system. Comput. Intell. **18**(2), 515–534 (2002)
12. Castelfranchi, C., Falcone, R.: Trust Theory: A Socio-Cognitive and Computational Model, vol. 18. Wiley, Hoboken (2010)
13. Coelho, H., da Rocha Costa, A.C.: On the intelligence of moral agency. Encontro Português de Inteligência Artificial, pp. 12–15 (2009)
14. Coelho, H., Trigo, P., da Rocha Costa, A.C.: On the operationality of moral-sense decision making. In: 2nd Brazilian Workshop on Social Simulation, pp. 15–20 (2010)
15. Cointe, N., Bonnet, G., Boissier, O.: Ethical judgment of agents' behaviors in multi-agent systems. In: AAMAS Proceedings, pp. 1106–1114 (2016)
16. Cointe, N., Bonnet, G., Boissier, O.: Multi-agent based ethical asset management. In: 1st Workshop on Ethics in the Design of Intelligent Agents, pp. 52–57 (2016)
17. Conte, R., Paolucci, M.: Reputation in Artificial Societies: Social Beliefs for Social Order, vol. 6. Springer Science & Business Media, Boston (2002)
18. Esfandiari, B., Chandrasekharan, S.: On how agents make friends: mechanisms for trust acquisition. In: Deception, Fraud, and Trust in Agent Societies Workshop, pp. 27–34 (2001)
19. Directorate-General for Economic and Financial Affairs: Impact of the current economic and financial crisis on potential output. Occasional Papers 49, European Commission, June 2009
20. Ganascia, J.-G.: Ethical system formalization using non-monotonic logics. In: 29th Annual Conference of the Cognitive Science Society, pp. 1013–1018 (2007)
21. Ganascia, J.-G.: Modelling ethical rules of lying with Answer Set Programming. Ethics Inf. Technol. **9**(1), 39–47 (2007)
22. Horsburgh, H.J.N.: The ethics of trust. Philos. Q. **10**(41), 343–354 (1960)
23. Johnson, R.: Kant's moral philosophy. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy, Summer edition, Metaphysics Research Lab, Stanford University (2014)
24. Josang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service proposition. Decis. Support. Syst. **43**(2), 618–644 (2007)
25. Lorini, E.: On the logical foundations of moral agency. In: 11th International Conference on Deontic Logic in Computer Science, pp. 108–122 (2012)
26. Marsh, S.: Formalising trust as a computational concept. Ph.D. Thesis, University of Stirling, Stirling (1994)
27. McIntyre, A.: Doctrine of double effect. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy, Winter edition, Metaphysics Research Lab, Stanford University (2014)
28. Muller, G., Vercouter, L., Boissier, O.: Towards a general definition of trust and its application to openness in MAS. In: 6th Workshop on Deception, Fraud and Trust in Agent Societies, pp. 49–56 (2003)
29. Ricoeur, P.: Oneself as Another. University of Chicago Press, Chicago (1995)

30. Rocha-Costa, A.: Moral systems of agent societies: some elements for their analysis and design. In: 1st Workshop on Ethics in the Design of Intelligent Agents, pp. 32–37 (2016)
31. Sabater, J., Sierra, C.: Review on computational trust and reputation models. Artif. Intell. **24**(1), 33–60 (2005)
32. Sabater-Mir, J., Vercouter, L.: Trust and reputation in multiagent Systems. Multiagent Systems, second Edition. Edited by Gerhard Weiss, MIT press. 381 (2013)
33. Saptawijaya, A., Moniz Pereira, L.: Towards modeling morality computationally with logic programming. In: Practical Aspects of Declarative Languages, pp. 104–119 (2014)
34. Timmons, M.: Moral Theory: An Introduction. Rowman & Littlefield, Lanham (2012)
35. Vercouter, L., Muller, G.: L.I.A.R.: achieving social control in open and decentralized multiagent systems. Appl. Artif. Intell. **24**(8), 723–768 (2010)
36. Yu, B., Singh, M.P.: Distributed reputation management for electronic commerce. Comput. Intel. **18**(4), 535–549 (2002)

Chapter 11
Rich and Poor: Simulating Social Policies to Build a Fairer Society

Diogo Costa, José Soares, and Pedro A. Santos

Abstract Inequality is present in many forms in most of today's societies at a scale that hinders their economic growth and the lives of the people who are part of it. We present an agent-based model that implements different economic redistribution policies and that can explain the emergence of its macroeconomic outcomes by means of the agents' interactions, as advocated by ACE. Ultimately, the simulation utilizes the generated data and applies several metrics to assess the inequality rate (e.g., Gini index). We were thus able to build a platform equipped with a series of features that allows the user not only to study the dynamics and the emergence of inequality in an economic system but also to analyze, compare, and measure the effect that different policy settings have in the inequality rate and other macroeconomic variables.

Keywords Agent-based simulations · RePAST · Income and wealth inequalities · Computational economics · ACE models

11.1 Introduction

The increase of inequality rates in the developed countries is giving rise to one of the main challenges of our time. Low economic growth and poor economic performance in the EU, as well as worldwide, have aggravated our concerns about wage disparity, skill and income inequality, and social exclusion [1, 3, 5, 10, 11, 23, 24]. Advanced economies have observed top income earners taking a larger share

D. Costa · J. Soares
Instituto Superior Técnico, University of Lisboa, Lisbon, Portugal

P. A. Santos (✉)
Instituto Superior Técnico, University of Lisboa, Lisbon, Portugal

INESC-ID Instituto de Engenharia de Sistemas e Computadores, Investigação e Desenvolvimento, Lisbon, Portugal
e-mail: pedro.santos@tecnico.ulisboa.pt

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_11

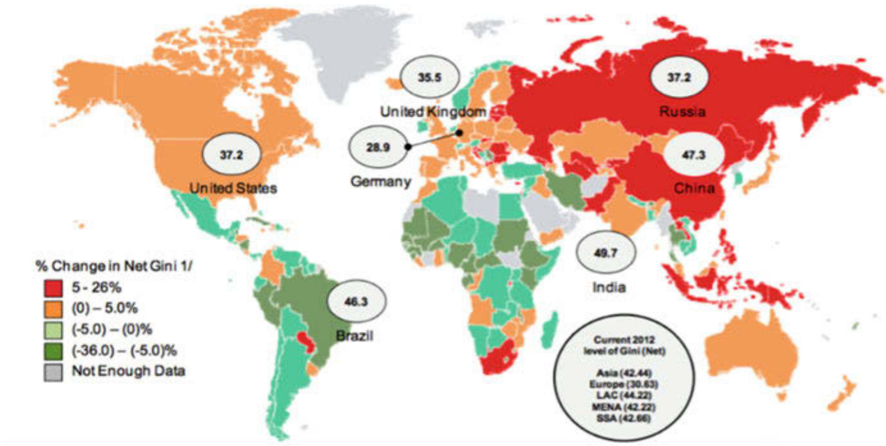


Fig. 11.1 Net income inequality variance from 1980 to 2012

of the total income generated by the economy [26, 40, 43]. In fact, the top 10% of income earners get nine times more income when compared to the bottom 10% [35].

In contrast, significant growth in emergent economies, along with redistribution efforts, lowered the income inequality [28, 46]. Figure 11.1 depicts changes in the distribution of economic disparities across the globe from 1980 to 2012.

The implications of economic inequalities are vast, from damaging trust and social cohesion to precluding economic growth, promoting macroeconomic instability, magnifying political influences, placing the decision power in the hands of few, neglecting the optimization of human resources, and increasing the risks of crisis, among many others [2, 4, 31, 39, 41]. In fact, there is a growing consensus that economic inequality leads to more fragile and less sustainable growth, while more equal economies sustain an ongoing growth [8, 38].

To combat economic inequalities, two of the main policy types proposed are policies to promote education among households and welfare policies through redistribution packages.

Regarding wealth redistribution policies, opinions often diverge: while Okun [37] argued that redistribution hurts growth due to lower labor and competition incentives, more recently Benabou [6, 7], Bleaney [9] and say that public expenditure in infrastructure, health, education, and social insurance provisions might be pro-growth.

Our objective is to present a model – Agent-based Model for Studying Inequality (AMoSI) – that permits to study and measure the impact of different types of tax revenue and redistribution strategies on inequality and other macroeconomic variables. The model is based on the principles of agent-based computational economics.

11.2 Agent-Based Computational Economics

Economies are complex adaptive systems, complex because there exist many constituents all interacting with each other and adaptive because they take actions in order to change the status quo or react to certain circumstances. In fact, the outcome of all the interactions among economic entities on the bottom emerges to the top and draws the overall performance of the economy. Mainstream frameworks of economic research such as DSGE models have taken a top-down approach to study economic behavior and analyze the effects of policy orientations. The latter have proven inadequate to cope with the recent turmoil in economic phenomena [21]. The agent-based theory provides a fresh and updated framework of analysis that minimizes the adversities from the previous methodologies for economic research. The ACE (agent-based computational economics) is the application of agent-based theory toward the study of economics.

ACE has been used to study various economic environments by constructing a relevant model in line with the requirements needed. It is still far from being a tool in the center of modern economic policy analysis due to both a lack of empirical model validation and robustness checks, but its importance has been increasing [15]. A comprehensive study about the origin of ACE and its types of agents is conducted by Chen et al. [12].

ACE applications are spread along a wide range of fields and themes. Malerba et al. [32] studied industrial policies and market designs with special regard for antitrust, entry-support, and public procurement policies and their consequences on the evolution of industry concentration and technology changes. Fagiolo and Tassier [21, 44] delivered new insights about the labor markets and its dynamics. Chen and Chie [13] used an agent-based model to capture the reason why taxa rates differ among countries with regard for social learning adaptive behavior in the lottery markets. Happe et al. [25] aimed at understanding the consequences of agriculture subsidies in order to provide potential orientation in policy design for the agriculture state of affairs within the EU. Oeffner [36] constructed an ACE model to study the impact of monetary policy from a central bank on its economy. Economic and fiscal policy studies have also been undertaken in [19, 20, 22, 34]. Mannaro et al. [33] created an agent-based model with the aim of reducing speculation in the financial markets by stabilizing foreign exchange and stock markets with Tobin tax-related policy designs. The financial markets were also the focus in [27, 29]. Ciarli et al. [14] aimed at studying the effects of changes in production and consumption patterns upon economic growth and income distribution. Dawid et al. [16, 17] created an ACE model to understand the effect of technology-oriented cohesion policies upon economic inequalities. Riccetti et al. [42] explored the effects of government intervention in the macroeconomic dynamics through unemployment benefits. At last, Tesfatsion and Lengnick [30, 45] generated a baseline example of an agent-based economic model, which results and reflects a few key important behaviors from macroeconomic phenomena. This was the starting line for the work we present.

11.3 The AMoSI Model

The model consists on a single-region, single-product decentralized market economy with two main co-existing sectors: households and consumer goods firms. Two fundamental markets bridge those sectors: the consumer goods market and the labor market. The consumer goods are non-perishable, which allows inventories to stock unsold products over time. Our baseline model considers no changes in individual productivities nor technology vintages and is similar to the one presented by Lengnick [30]. The full model includes tax collection and redistribution policies, as well as the effects of education. Each iteration of the simulation represents an economic cycle with the duration of 1 month.

AMoSI runs on an iteration loop that integrates three agent classes. The classes of the agents can be divided in *households, firms*, and *government*. Each of these agents interacts with one another to produce the economy's macroeconomic outputs (inflation, GDP, employment rates, etc.). In order to produce these outputs, agents make use of a labor market and a goods market.

Figure 11.2 illustrates AMoSI's economic cycle. The cycle comprehends seven stages that summarize all the agents' interactions with the environment. All parameters' descriptions and values used for the agents' decisions can be found in Table 11.1.

During the planning stage, the amount of goods a firm, f, aims to produce in a simulation step, t, depends on the amount of goods that it sold in $t-1$, compared to the number of goods that it stocked. Thus, we establish that if a firm was able to sell more than a given percentage of what it stocked, γ_f, then it is considered safe to increase the production relative to $t-1$; but if the sales fall below the threshold defined by δ_f, then the firm should decrease its production levels to save costs. In other words, if $\text{soldGoods}_{f,\,t-1} \geq \text{stock}_{f,\,t-1} \cdot \gamma_f$, then firms will determine their target production according to the following equation:

$$\text{targetProduction}_{f,t} = \text{production}_{f,t-1} \cdot (1+\beta) \tag{11.1}$$

Contrarily, if $\text{soldGoods}_{f,\,t-1} \leq \text{stock}_{f,\,t-1} \cdot \delta_f$, then the following equation applies:

$$\text{targetProduction}_{f,t} = \text{production}_{f,t-1} \cdot (1-\beta) \tag{11.2}$$

The parameters γ_f and δ_f are heterogeneous among firms, i.e., each firm has their own upper and lower thresholds to decide when to increase or decrease their production levels. On the other hand, β is a constant parameter that corresponds to the production's increase/decrease percentage for all firms.

The job application stage in Fig. 11.2. is concurrent with the planning stage and simply represents the ability of an unemployed household agent to join the labor market, so that he receives job offers from the firms.

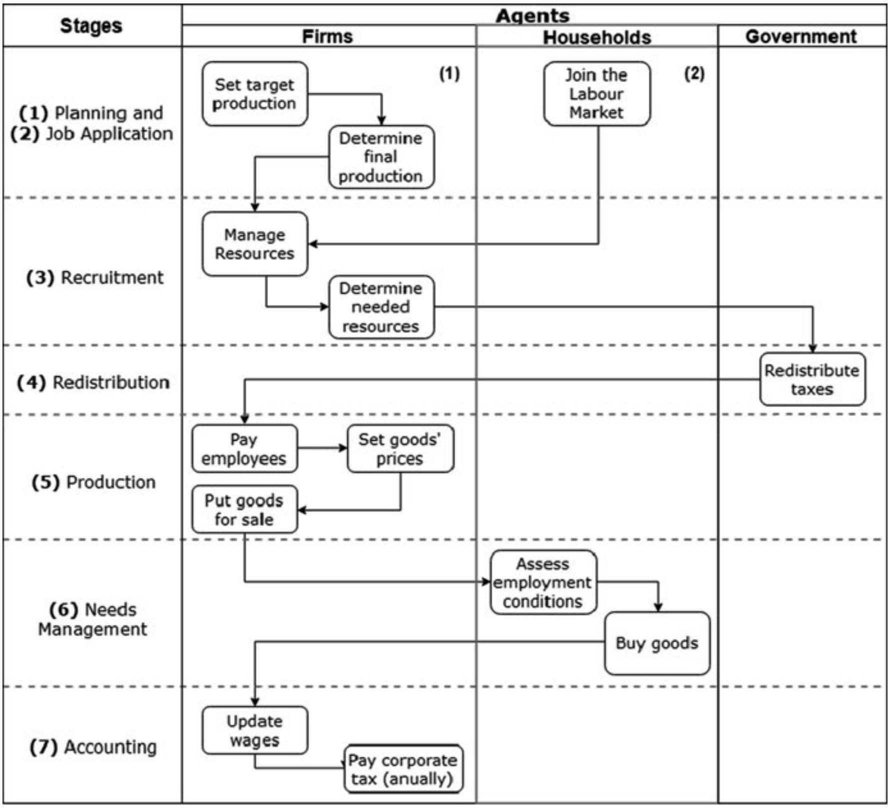


Fig. 11.2 AMoSI's Stages

Household agents are assigned with one of three education levels, eduTier_h, which dictate their labor productivity, λ_h, and are the inputs for the firms' production processes. The higher the education level, the faster the learning speed, and consequently, it will result in higher productivities. The following equation displays the explained behavior:

$$\lambda_{h,t} = \lambda_{h,t-1} + \frac{\text{eduTier}_h}{\lambda_{h,t-1}} \tag{11.3}$$

On the other hand, an unemployed household loses its individual skills at a constant rate, by a factor *prodDecrease*, which means that during the time it is unemployed, it will lose productivity. As such, in every tick, $\lambda_{h,t} \leftarrow \lambda_{h,t-1} \cdot (1 - \text{prodD}\textit{ecrease})$.

This characteristic of the households plays a crucial role in the firms' recruitment stage. Having employees whose productivity is defined by their education implies that the firm's production is constrained by the quantity and quality of labor available in the labor market, besides the budget that they can dispose of in hiring.

Table 11.1 Initial parameter settings for the post-baseline model

Notation	Definition	Value
H	Number of households	500
F	Number of firms	50
inicial_wealth_h	Initial households' funds (at $T = 0$)	0
inicial_wealth_f	Initial firms' funds (at $T = 0$)	250
inicial_wealth_g	Initial funds from government (at $T = 0$)	0
r_h	Utility reduction factor by household h	$[0,4;0,6]$
initial_resWage_h	Reservation wage of household h	3
φ	Reservation wage reduction factor	$[0,05;0,2]$
$\text{initial\_targetProd}_f$	Initial target production by firm f	10
γ	Stock lower limit for production increases	$[0,85;0,95]$
δ	Stock upper limit for production decreases	$[0,45;0,55]$
β	Production increase/decrease factor	0,05
$\text{initial\_price}_{b,f}$	Price of the initial batch of goods by firm f	3
σ	Stock lower limit for price increases	0,9
μ	Stock upper limit for price decreases	0,8
θ	Price increase/decrease factor	$[0,08;1,2]$
λ	Households' productivity	2
$\text{initial\_wageOffer}_f$	Initial proposed wage by firm f	5
ε	Wage offer increase/decrease factor	$[0,08;0,15]$
ϑ	Visibility constraint in the goods market	0,7
η	Visibility constraint in the labor market.	1
$\text{MinW}age$	Minimum wage for every Household.	1
ν	Percentage of the last wage paid in unemployment benefits	1
ϖ	Maximum unemployment time under unemployment benefits (months)	9
eduTier_h	Productivity associated with respective education level from household h	$\{1,75;1,45;1,2\}$
prodDecrease	Productivity decrease factor for unemployed households	0,1
τ	Stock goods discount	0,6

Furthermore, household agents accept or refuse a job offer based on their reservation wage ($\text{resWage}_{t,\ h}$), which represents the minimum amount of money they are willing to work for. Thus, if a firm's wage proposal is greater than the reservation wage, the household will accept the offer; otherwise it will not. This behavior, combined with the different productivity levels and available resources and budget, means that the *targetProduction* may not always be achieved and it will have to be updated according to the number of hired resources and what they can produce.

The redistribution stage is represented by the government, who implements three tax concepts, the value-added tax (VAT) – applied to the goods, the corporate tax (CT) – applied to the firms' profits, and the income tax – applied to the households' salaries, and three benefits: earned income tax credit (EITC), unemployment benefit, and minimum benefit. While the taxes are exogenously parameterized before the simulation starts, the benefits are applied according to a specific algorithm.

Whenever a household is unemployed for less than a given number of simulation steps, ϖ, it is entitled to the unemployment benefit. This benefit consists in a percentage, ν, of its last wage. However, if the household agent is unemployed for longer than ϖ, then the minimum benefit will be applied, which corresponds directly to the minimum wage (after taxes). In case the household is employed, a work incentive will be applied. The EITC will benefit those households according to their income distribution; the greater their salary, the smaller the bonus.

In the production phase, when determining the goods' price, firms must take into consideration the unitary costs of production – which correspond directly to the sum of all employees' salary – so that they shall not incur in debt.

The pricing decisions of a firm's goods are analogous to Eqs. (11.1 and 11.2), only in this case, it is not the thresholds (σ and μ, see Table 11.1) that vary from firm to firm. Instead, the percentage of increase/decrease in their goods' prices will vary with a heterogeneous factor θ_f. Additionally, firms perform a discount, τ, on all products that are kept in inventory for more than a year (12 simulation steps) to liberate stock and attenuate losses.

Still during the production phase, the way firms decide on the salary they offer to new employees depends on whether they successfully hired the desired number of resources. If so, the offers decrease by a heterogeneous factor, ε_f, which means there is a high supply; otherwise the offers increase by the same factor, which means there is a high demand. Additionally, all employees' salaries are increased by a small amount when the employment rates of the economic system surpass 95%.

The needs management stage encompasses the households' consumption and a *self-assessment* on their current working conditions. The consumption patterns are heterogeneous among these agents, which means that they have different inherent propensities to consume (and therefore to save). Let j be the j^{th} good to be consumed and $r \in \,]0, 1[$ a heterogeneous parameter (see Table 11.1). For each household h:

$$\text{utility}_h(j) = \text{disposableIncome}_{h,t} \cdot {r_h}^{j-1}, \quad 0 < r < 1 \tag{11.4}$$

Equation (11.4) means that the marginal utility of a good is decreasing as j increases. The household's purchase decision is then carried out by measuring the good's utility against its price. The disposable income of a household in t corresponds to the sum of the savings from the previous simulation step: its net wage plus its employment benefits.

Additionally, households try to adapt their employment conditions by adjusting their reservation wage, introduced earlier. Thus, similarly to the households' unemployment benefits, if it is unemployed for less than a given number of simulation steps, ϖ, then they will try to update their reservation wage to a percentage, ν, of their last wage. However, every tick, they will keep on decreasing their reservation wage by a heterogeneous factor, φ_h, so that their chances of getting a job increase. This corresponds to a household lowering their salary expectations and accepting more poorly paid jobs. Lastly, if they have been unemployed for longer than ϖ, then they will not consider their last wage and will simply keep on decreasing their expectations until they reach the minimum wage of the economic model.

On a last note, the goods market is constrained by asymmetric information: the households are not aware of all the firm agents in the system and can only attempt to buy the cheapest goods available in their own network of "visible" firms. Then, the household agent buys a good j if the utility of j (see Eq. (11.4)) is greater than the value of its market price; otherwise the agent will find that the good is not worth buying.

11.4 Results Validation

The model was implemented using a Java toolkit called RePAST,[1] which allowed us to describe the state and evolution of the simulated economies. For the present paper, two scenarios were chosen to demonstrate the features that AMoSI provides to study anti-inequality policies.

The first scenario *(S0)* consists of an economy free from any kind welfare redistribution system, which means there are no taxes nor benefits to be distributed by the government. This is the baseline model, with only education effects added.

The second scenario *(SR)* represents an economy where there is an active redistribution system, with VAT, IRS, and CT at a rate of 10%, 40%, and 15%, respectively. Additionally, all the three existing benefits available in the AMoSI model are also distributed.

In both scenarios, the population of households is divided into three tiers of education, where the first, second, and third tiers consist of 5%, 15%, and 80% of the number of household agents, respectively. Each simulation was run 100 times, and the results started counting from the 1200th tick, which consists of the simulation's *warm-up* time. Table 11.1 comprises the initial set of parameters chosen to initiate the simulations.

The AMoSI's results regarding both scenarios were validated by checking the existence of *endogenous business cycles* and the *Phillips* and *Beveridge curves* [18].

11.4.1 Endogenous Business Cycles

The first validation check used in this process is the model's ability to independently generate endogenous business cycles. Figures 11.3 and 11.4 present the wealth distribution among the various agents in both scenarios, S0 and SR, respectively, for a period of 50 years (600 simulation steps after *warm-up*).

[1] http://repast.sourceforge.net/docs/RepastReference.pdf

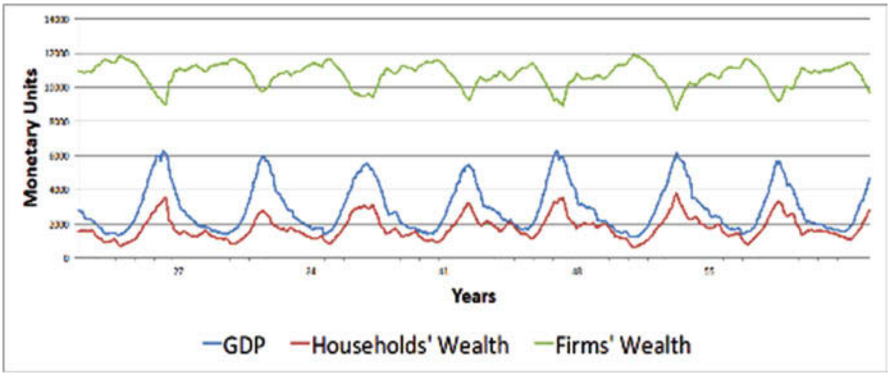


Fig. 11.3 GDP, households', and firms' wealth for the S0 scenario

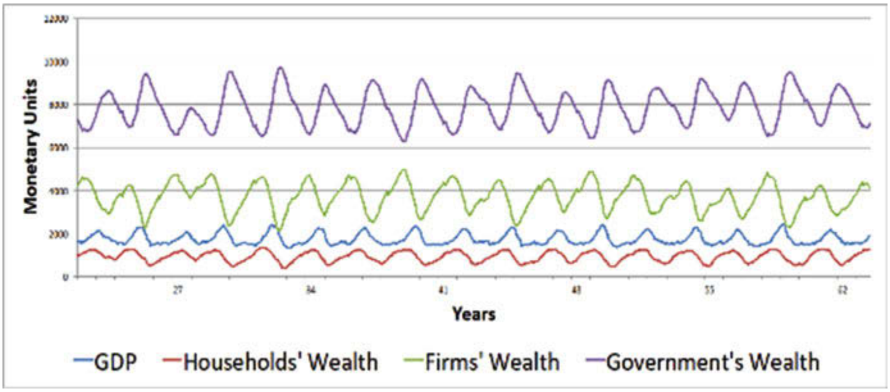


Fig. 11.4 GDP, households', firms', and government's wealth for the SR scenario

From the last two figures, it becomes clear that GDP and the households' wealth are positively correlated, while GDP and the firms' wealth are negatively correlated. The GDP fluctuates over periods of expansion and depression in line with the households' purchasing power at a given moment, as more wealth in the household's side results in a growing demand and therefore higher GDP.

The GDP is a function of four main variables: goods' prices, aggregate consumption, wages, and employment rates. Each of these variables has a different impact on the final GDP. The next two figures attempt to deliver a few important insights on the influence of these variables on the GDP outlook and the generated business cycles. The graphics are only with respect to the S0 scenario, but the behavior is analogous to the SR scenario.

The graphics below allow us to draw conclusions regarding the emergence of business cycles generated by the model. When both employment (and consequently the production output) and consumption reach their peak, prices are still increasing because there is no constraint on pricing as opposed to the total number of

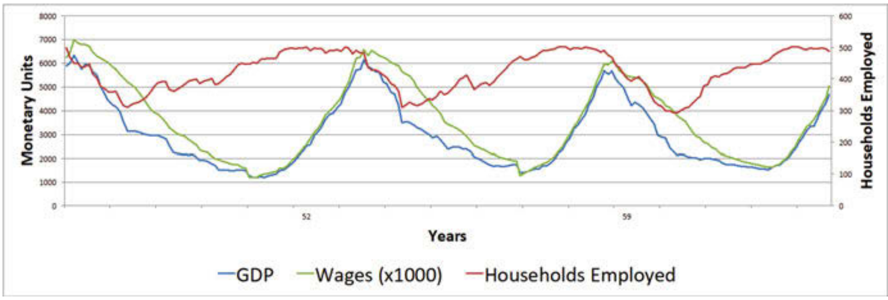


Fig. 11.5 GDP, number of households employed, and average wages for S0

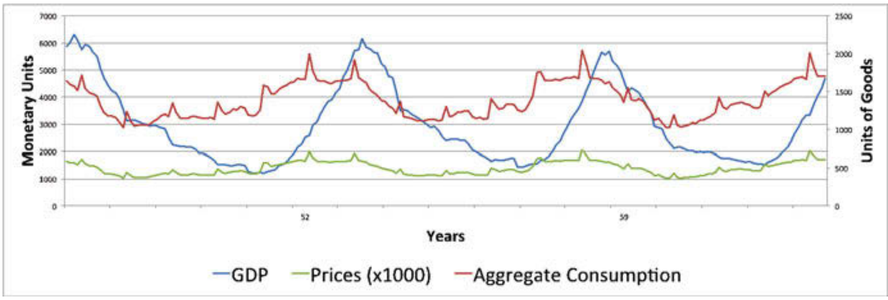


Fig. 11.6 GDP, aggregate consumption, and average prices for S0

households dictating the maximum production output of the economy. Therefore, if consumption justifies inflation, prices increase. Wages also increase because of high employment rates; since it gets more difficult to find resources, firms increase their wage offers. These two phenomena reach their limit when prices are unsustainable, and households do not find enough utility in purchasing goods at such high values, which takes the economy to the business cycle's turning point (Figs. 11.5 and 11.6).

The opposite occurs in the negative slope of the cycle. GDP is highly influenced by inflation and deflation, which are driven by consumption. Therefore, even though real GDP is often the most used metric to compare different economies, it fails to measure the general welfare of the population when there is a rapid variation of prices and salaries.

11.4.2 Phillips and Beveridge Curves

To complete the validation process, we had to assess whether the model complies with both the Phillips and the Beveridge curves [18].

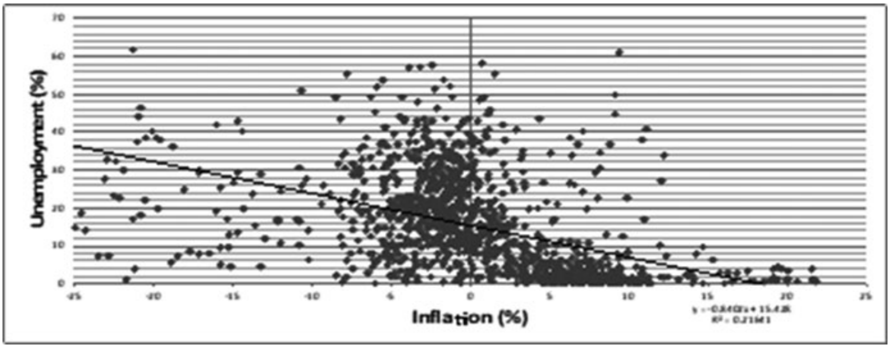


Fig. 11.7 Phillips curve for S0

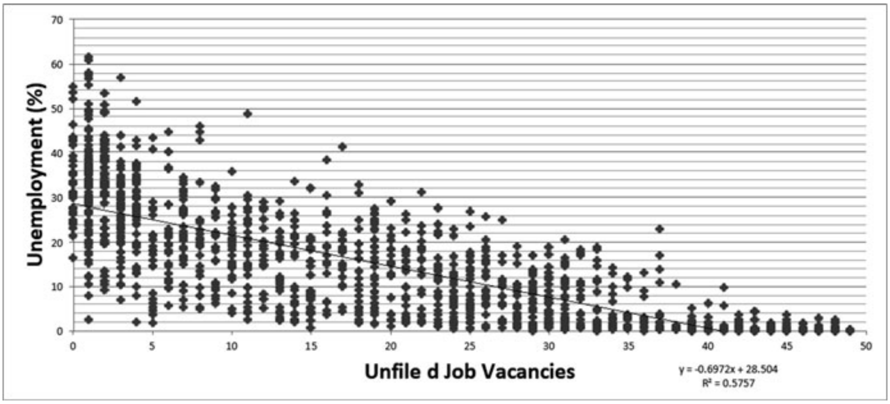


Fig. 11.8 Beveridge curve for S0

The Phillips curve is an empirical economic relation that states an inverse correlation between the unemployment rate and inflation, which means that unemployment decreases as inflation increases.

The Beveridge curve is the graphical representation of the relationship between the number job vacancies and the unemployment rate. According to the Beveridge curve, there is a negative feedback between the unemployment rate and the number of job vacancies, that is, when the unemployment increases, the number of vacancies decreases.

The following four figures present the graphical representations of both these empirical laws in both scenarios under assessment (Figs. 11.7, 11.8, 11.9 and 11.10).

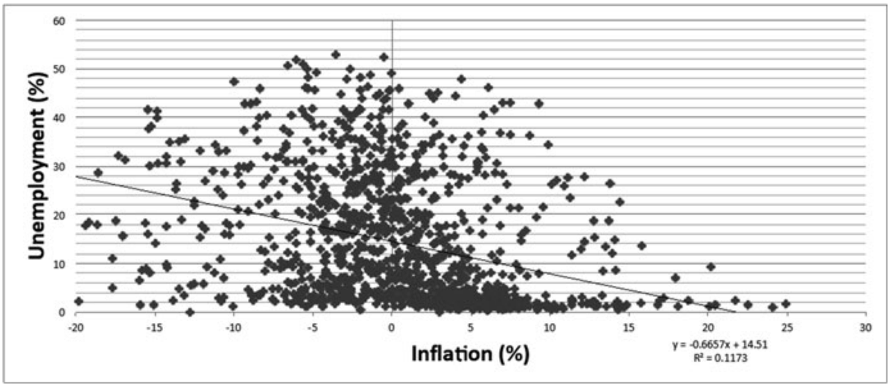


Fig. 11.9 Phillips curve for SR

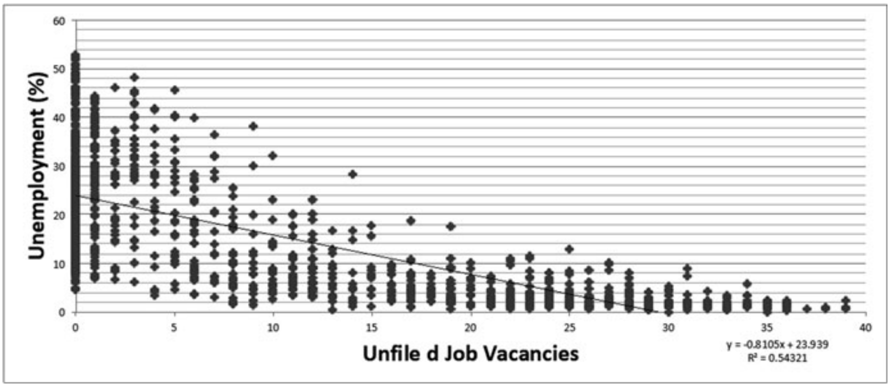


Fig. 11.10 Beveridge curve for SR

11.5 Economic Inequality Analysis

As previously mentioned, the main goal of the AMoSI model is to create a platform to study normative understanding and policy analysis upon economic inequalities. In this section, the core inequality dynamics are presented, and further ahead, both the S0 and SR scenarios are compared in order to assess whether the welfare redistribution policy is, in fact, beneficial for this artificial society.

Understanding how income inequality emerges in a society is one of the primary goals of the AMoSI model. The following graphics are with respect to the S0 scenario (Figs. 11.11 and 11.12).

From both the previous figures, it is possible to observe a positive correlation between equality and consumption. In fact, higher consumption levels lead to higher employment, and employment is highly correlated to economic equality with almost 90% *goodness-of-fit*. In this scenario, unemployed households do not have any

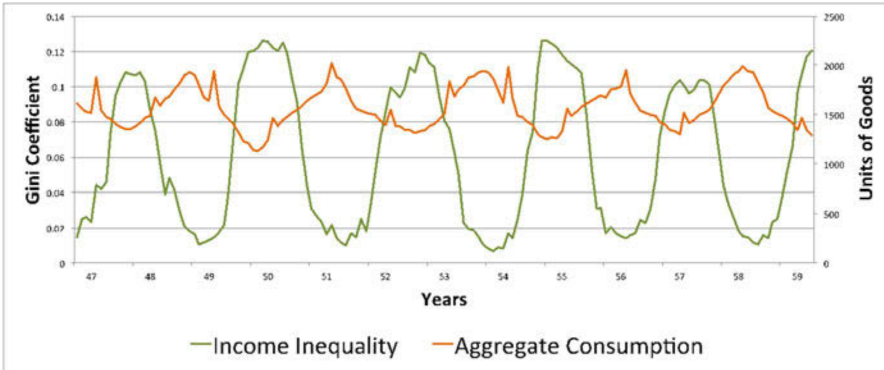


Fig. 11.11 Income inequality and aggregate consumption for S0 scenario

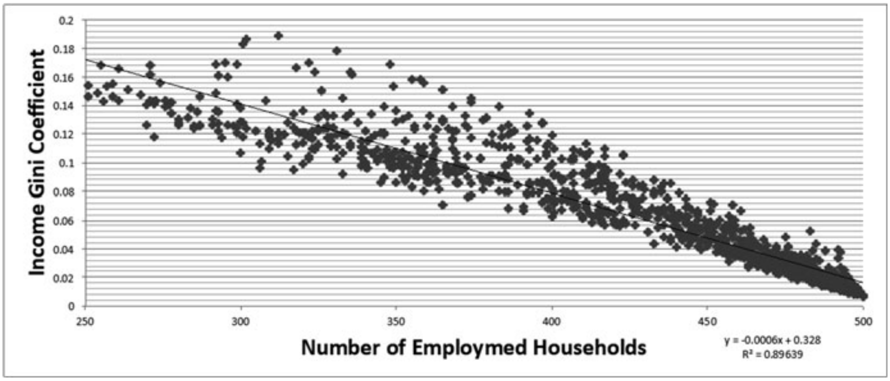


Fig. 11.12 Income inequality and employment correlation for S0 scenario

source of income, which prevents them from purchasing their desired quantities of consumer goods. Contrarily, the employed households are the only ones with enough purchasing power to consume. The gap between unemployed and employed households' income only starts to diminish when aggregate consumption increases, which forces the aggregate production to increase and, consequently, so does the employment rate.

These dynamics change when other sources of income are added to the economic system or when households have different incentives to work. In the SR scenario, households are entitled with a minimum wage and an unemployment benefit if unemployed, which allows them to purchase consumer goods even when they do not get a salary. Also, households get a bonus from the government for being employed, if their salary is low, which enhances the purchasing power of employed agents and functions as an incentive to work. Firms also benefit from the EITC benefit because it becomes easier for them to hire new households, since the responsibility

Table 11.2 Average aggregate consumption and income inequality for S0 and SR scenarios

	S0	SR
Average aggregate consumption	1486	**1532**
Average income inequality	0,244	**0,064**

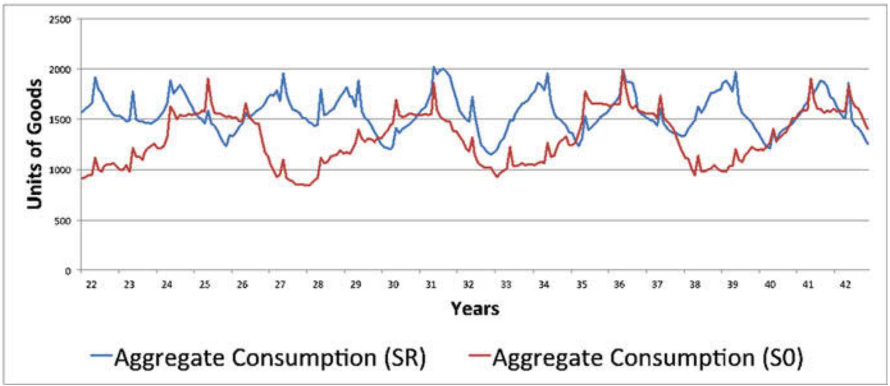


Fig. 11.13 Aggregate consumption for the SR and S0 scenarios

of satisfying the households' reservation wage is no longer exclusive of the firms but also an indirect responsibility of the government.

As previously mentioned, GDP is largely used to compare different economies, but it seems to be too sensitive to inflation and deflation phenomena, which will not account for the welfare of the population. On the other hand, aligning aggregate consumption with income inequality may provide a much better understanding on the population's general welfare and equality.

Table 11.2 comprises the average of the two metrics previously mentioned. Also, the following two figures present both scenarios under study regarding its aggregate consumption and income inequality dynamics (Fig. 11.13).

From Table 11.2, it is possible to conclude that, in average, households consume more in the SR scenario when compared to the S0 scenario. At the same time, the SR scenario results in a more equal society as shown by the lower income Gini coefficient. In the S0 scenario, income and wage overlap because wages are the only income source of the households, while in SR the practice of the government distributing three types of benefits completely changes this behavior. In fact, from Fig. 11.14., it is possible to observe the gap between wage and income inequality for the SR scenario, which is due to the presence of the welfare redistribution policy.

From the previous graphics and table, it is possible to demonstrate that a simple economy, such as the one implemented by the AMoSI model, displays a better performance when there is an active welfare redistribution policy with three taxes and benefits included. The AMoSI model did not only allow us to assess the simulation's outcomes but also to formulate other insights on how economic inequalities emerge in a modern society. This aspect is most beneficial for policymakers to tackle this issue in the most adequate way possible.

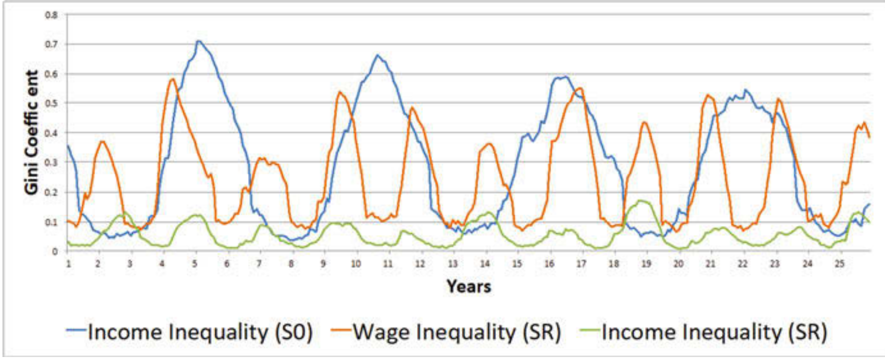


Fig. 11.14 Income and wage inequality for the SR and S0 scenarios

11.6 Conclusions

Economic inequalities have been in the center of the economists' attention throughout the years. They threaten social and economic cohesion while also reducing growth. The present work introduced a new approach to the study of economic policies and aimed to find a means to tackle this issue. As such, a multi-agent simulation composed by households, firms, and a government was developed in order to simulate an economic model and assess the impact of different economic policies on the system, giving birth to the AMoSI.

This model relied on the traditional welfare redistribution policy to assess whether these policies contribute toward a more reasonable and equal society while at the same time not precluding the system's economic performance.

In order to perform this assessment, two metrics were chosen to help guide our analysis: the households' aggregate consumption and the Gini coefficient of their income. Throughout the simulation runs, the system showed consistently higher levels of aggregate consumption as the income Gini coefficient decreased, meaning that the higher the consumption levels, the lower the inequality rate of the system tends to be.

Moreover, the AMoSI allows the user to study the microeconomic dynamics behind the emergence of inequality, which can pose as a major benefit for policymakers trying to create policy sets to attenuate the inequality rates. Therefore, the AMoSI is successful in both creating new insights on the dynamics of inequalities and also allowing the analysis of specific economic policies. In fact, it is possible to use our model to study both the impact and effectiveness of individual taxes and benefit levels as well as their interaction effect on growth and inequality reduction.

On the other hand, there are still some points to improve. The AMoSI is still a quite simplistic model of a modern economy, having its focus on the firms' and households' behavior and interactions. In fact, introducing other products, a financial system, and active government constraints regarding the public budget will add a significant amount of complexity to the model. However, by doing so, it may converge into an even more realistic version of an economic system. Nevertheless, our model does pose as a valid initial framework for the analysis of economic policies.

Acknowledgments This work was supported by national funds through the Fundação para a Ciência e a Tecnologia (FCT) with reference UID/CEC/50021/2019.

References

1. Acemoglu, D.: Why do new technologies complement skills? Directed technical change and wage inequality. Q. J. Econ. **113**(4), 1055–1089 (1998)
2. Alesina, A., Perotti, R.: Income distribution, political instability and investment. Eur. Econ. Rev. **40**(6), 1203–1228 (1996)
3. Atkinson, A.B., Piketty, T., Saez, E.: Top incomes in the long run of history. J. Econ. Lit. **49**(1), 3–71 (2011)
4. Bardhan, P.K.: Scarcity, Conflicts, and Cooperation: Essays in the Political and Institutional Economics of Development. MIT Press, Cambridge (2005)
5. Becker, G.S., Chiswick, B.R.: Education and the distribution of earnings. Am. Econ. Rev. **56**(1/2), 358–369 (1966)
6. Benabou, R.: Unequal societies: income distribution and the social contract. Am. Econ. Rev. **90**(1), 96–129 (2000)
7. Benabou, R.: Tax and education policy in a heterogeneous-agent economy: what levels of redistribution maximize growth and efficiency? Econometrica. **70**(2), 481–517 (2002)
8. Berg, A., Ostry, J.D.: Inequality and Unsustainable Growth: Two Sides of the Same 28 Coin?. IMF Staff Discussion Note 11/08. International Monetary Fund, Washington (2011)
9. Bleaney, M., Gemmell, N., Kneller, R.: Testing the endogenous growth model: public expenditure, taxation, and growth over the long run. Can. J. Econ. **34**(1), 36–57 (2001)
10. Brandolini, A., Rosolia, A., & Torrini, R. (2010). The Distribution of Employees' Labour Earnings in the EU: Data, Concepts and First Results. Eurostat Working Paper, European Union
11. Card, D., DiNardo, J.E.: Skill biased technological change and rising wage inequality: some problems and puzzles. NBER Working Paper 8769, National Bureau of Economic Research, Cambridge, (2002)
12. Chen, S.: Varieties of agents in agent-based computational economics: a historical and an interdisciplinary perspective. J. Econ. Dyn. Control. **36**(1), 1–25 (2012)
13. Chen, S., Chie, B., Lee, C.: Agent-Based Modeling of Lottery Markets with the Expected-Utility Paradigm. 2005 IEEE Congress on Evolutionary Computation (2005)
14. Ciarli, T., Lorentz, A., Savona, M., Valente, M.: Structural change of production and consumption: a micro to macro approach to economic growth and income distribution. Metroeconomica. **61**(1), 180–218 (2008)
15. Crivellaro, E.: College wage premium over time: trends in Europe in the last 15 years. University of Venice, Working Paper 03/2014 (2014)
16. Dawid, H., Fagiolo, G.: Agent-based models for economic policy design: introduction to the special issue. J. Econ. Behav. Organ. **67**(2), 351–354 (2008)
17. Dawid, H., Harting, P., Neugart, M.: Cohesion policy and inequality dynamics: insights from heterogeneous agents macroeconomic model. Bielefeld Working Papers in Economics and Management, 26 (2013)
18. Delli Gatti, D., Gaffeo, E., Gallegati, M., Giulioni, G., Palestrini, A.: Emergent Macroeconomics. Springer, Milan (2008)

19. Dosi, G., Fagiolo, G., Napoletano, M., Roventini, A., Treibich, T.: Fiscal and monetary policies in complex evolving economies. J. Econ. Dyn. Control. **52**(2014/22), 166–189 (2015)
20. Dosi, G., Fagiolo, G., Roventini, A.: An evolutionary model of endogenous business cycles. Comput. Econ. **27**(1), 3–34 (2006)
21. Fagiolo, G., Dosi, G., Gabriele, R.: Towards an Evolutionary Interpretation of 92 Aggregate Labor Market Regularities. Sant'Anna School of Advanced Studies, Pisa (2004)
22. Fagiolo, G., Roventini, A.: Macroeconomic policy in DSGE and agent-based models Redux: new developments and challenges ahead. J. Artif. Soc. Soc. Simul. **20**(1), 1 (2017)
23. Feenstra, R.C., Hanson, G.H.: Globalization, outsourcing, and wage inequality. Am. Econ. Rev. **86**, 240–245 (1996)
24. Frederiksen, A., Poulsen, O.: Increasing income inequality: productivity, bargaining and skill-upgrading. IZA Discussion Papers 4791, Institute for the Study of Labor, Bonn, Germany (2010)
25. Happe, K., Kellermann, K., Balmann, A.: Agent-based analysis of agricultural policies: an illustration of the agricultural policy simulator AgriPoliS, its adaptation and behavior. Ecol. Soc. **11**, 49 (2006)
26. Kierzenkowski, R., Koske, I.: Less income inequality and more growth – Are they compatible? Part 8. The drivers of labour income inequality – A literature review. OECD Economics Department Working Papers 931 (2012)
27. Kyrtsou, C., Malliaris, A.G.: The impact of information signals on market prices when agents have non-linear trading rules. Econ. Model. **26**, 167–176 (2009)
28. Lakner, C., Milanovic, B.: Global income distribution: from the fall of the Berlin Wall to the Great Recession. WB Policy Research Working Paper 6719, World Bank, Washington (2013)
29. Lebaron, B.: Agent-based computational finance: suggested readings and early research. J. Econ. Dyn. Control. **24**(5–7), 679–702 (2000)
30. Lengnick, M.: Agent-based macroeconomics: a baseline model. J. Econ. Behav. Organ. **86**(2013), 102–120 (2013)
31. Lichbach, M.I.: An evaluation of 'does economic inequality breed political conflict?' Studies. World Polit. **41**(4), 431–470 (1989)
32. Malerba, F., Orsenigo, L.: Innovation and market structure in the dynamics of the pharmaceutical industry and biotechnology: towards a history-friendly model. Ind. Corp. Chang. **11**, 667–703 (2002)
33. Mannaro, K., Marchesi, M., Setzu, A.: Using an artificial financial market for assessing the impact of Tobin-like transaction taxes. J. Econ. Behav. Organ. **67**(2), 445–462 (2008)
34. Neugart, M.: Labor market policy evaluation with ACE. J. Econ. Behav. Org. Elsevier **67**(2), 418 (2008)
35. OECD: Income Inequality Update - June 2014 (2014). https://www.oecd.org/els/soc/OECD2014-Income-Inequality-Update.pdf
36. Oeffner, M.: Agent–based keynesian macroeconomics — an evolutionary model embedded in an agent–based computer simulation. University of Wurzburg, Faculty of Economics and Business Administration, (2008)
37. Okun, A.M.: Equality and Efficiency: the Big Trade-Off. Brookings Institution Press, Washington (1975)
38. Ostry, J.D., Berg, A., Tsangarides, C.: Redistribution, Inequality, and Growth. IMF Staff Discussion Note 14/02, International Monetary Fund, Washington (2014)
39. Piketty, T.: Capital in the Twenty-First Century. Harvard University Press, Cambridge, MA (2014)
40. Piketty, T., Saez, E.: Income inequality in the United States, 1913–1998. Q. J. Econ. **118**(1), 1–39 (2003)
41. Rodrik, D.: Where did all the growth go? External shocks, social conflict, and Growth Collapses. J. Econ. Growth. **4**(4), 385–412 (1999)
42. Riccetti, L., Russo, A., Gallegati, M.: Unemployment benefits and financial leverage in an agent based macroeconomic model. Economics. **7**(2013–42), 1–44 (2013)

43. Saez, E., Zucman, G.: Wealth inequality in the United States since 1913: evidence from capitalized income tax data. Quart. J. Econ. **131**, 519–578 (2016)
44. Tassier, T.: Social network structure, segregation, and equality in a labor market with referral hiring. J. Econ. Behav. Org. **66**(3–4), 514–528 (2005)
45. Tesfatsion, L.: Chapter 16 agent-based computational economics: a constructive approach to economic theory. In: Handbook of Computational Economics, pp. 831–880, North-Holland (2006)
46. Tsounta, E., Osueke, A.: What Is Behind Latin America's Declining Income Inequality? IMF Working Paper 14/124, International Monetary Fund (2014)

Chapter 12
Multi-scale Validation of an Agent-Based Housing Market Model

Koen de Koning and Tatiana Filatova

Abstract Validation is a vital step in any model that aspires to have an impact. In agent-based computational economics in particular, it is essential because, unlike most neoclassical economic models, the models are vastly complex and often unique in their kind [1, 2]. Importantly, they often require disaggregated data to specify attributes, behavioural rules and interactions among agents. Therefore, agent-based models require more time and effort to validate. There are numerous ways of validation in agent-based modelling discussed in the literature, e.g [1, 3–7]. Reviewing them is beyond the scope of this paper. Here we will focus on two dimensions of validation: empirical versus theoretical validation and micro- versus macroscale validation. In our agent-based model of the housing market, we have applied both ends of the two spectra.

12.1 Introduction

Validation is a vital step in any model that aspires to have an impact. In agent-based computational economics in particular, it is essential because, unlike most neoclassical economic models, the models are vastly complex and often unique in their kind [1, 2]. Importantly, they often require disaggregated data to specify attributes, behavioural rules and interactions among agents. Therefore, agent-based models require more time and effort to validate. There are numerous ways of validation in agent-based modelling discussed in the literature, e.g [1, 3–7]. Reviewing them is beyond the scope of this paper. Here we will focus on two dimensions of validation: empirical versus theoretical validation and micro- versus macroscale validation. In our agent-based model of the housing market, we have applied both ends of the two spectra.

K. de Koning (✉) · T. Filatova
Department of Governance and Technology for Sustainability (CSTM), University of Twente, Enschede, The Netherlands
e-mail: k.dekoning@utwente.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_12

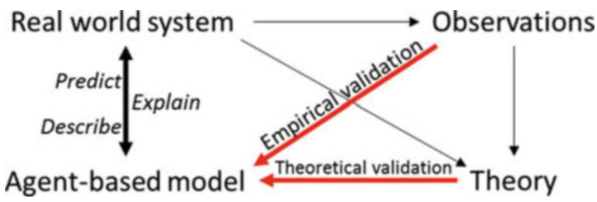


Fig. 12.1 Validation process for agent-based models. Empirical validation is directly based on observations of the real-world system, while theoretical validation – like the name suggests – is done by making use of theories, potentially derived from observations and our understanding of the real-world system or confirmed by them. Note that the arrows from observations to theory (induction) are also empirical validation but are beyond the focus of this paper

Empirical versus theoretical validation can be presented as a trade-off between making the model as detailed and realistic as possible – demanding empirical validation – and as simple and stylised as possible, which mostly demands theoretical validation. These two options of validating agent-based models are not mutually exclusive. Yet, the trade-off between detailed representation of a particular case study and generalisability governs where the modeller should put most of her effort. Moreover, she should realise that theoretical validation is often a lot less costly and time-consuming than empirical validation. Figure 12.1 summarises our conceptual overview of these two choices in an agent-based model where we try to describe, explain, explore as well as predict real-world phenomena. Processes in these real-world systems can be observed and studied. A modeller can choose to spend her time collecting data and formulating stylised facts based on these observations and use that as input for the agent-based model – empirical validation. When time is scarce, however, or when a modeller wants to build a model that is more stylised and generalisable, she will need to rely on theories – theoretical validation. Ideally, these behavioural theories should be passing an empirical test independently of a particular agent-based exercise, as many theories in psychology do [8, 9].

In this paper, we discuss our experience with an agent-based model of the housing market that is validated both empirically and theoretically across multiple scales. In our model, we attempt to seek a balance in scale and realism by combining empirical and theoretical validation methods and applying them on multiple scales (agent-level and market-level). We highlight what we learned from this example and how it is useful for other modellers, especially in environmental applications of agent-based computational economics where both spatial and behavioural data are to be combined.

12.2 Model Description

We built an agent-based model to explore how urban housing markets evolve in the presence of climate-driven floods and behavioural biases on the agent level. Climate change imposes severe consequences for urban development in hazard-prone coastal and delta areas. The issue is complicated by the fact that disaster risk is spatially correlated with rich environmental and urban amenities of those locations. They drive population clustering and growth of property values in hazard-prone areas, rapidly increasing exposure and vulnerability. Individual preferences for locations play a crucial role in the formation of spatial patterns in urban land markets. Yet, increasing climate-related risks may alter choices of individual economic agents, which enforces the potential for critical transitions from the bottom-up. In complex systems, changes in individual expectations driven by new information and emotions could lead to major abrupt shifts in the aggregated market dynamics.

We address the methodological challenge of studying non-marginal abrupt land use transitions by incorporating adaptive expectations about land market dynamics into a spatial agent-based model. It is a complete model of the property market in both a coastal and an inland city that experience regular floods. Extensive datasets of both GIS and housing transaction data are employed [10, 11] in addition to income and housing budget data from the USA [12–14]. Our current model is based on an existing model of a flood-prone housing market, the background and ODD+D description for which can be found elsewhere [15]. The current version makes significant improvements on the price expectations' procedure [16] and the structure of the utility function. In the model, we simulate household decisions in the flood-prone housing market, in particular, (1) households deciding on whether or not to sell their property, (2) a realtor agent advising sellers on their ask price, (3) buyers looking for properties within their budget, (4) buyers comparing properties based on expected utility calculations and deciding to bid on a property, (5) sellers entering negotiation with buyers that offer the highest bid and (6) buyers and sellers negotiating over prices. Flood risk enters the calculations in steps 2 and 4. The capitalised discount for properties at risk of flood is calculated in step 2, and individual risk perception enters the utility calculation in step 4. The utility calculation is based on buyers' housing budgets, preferences for neighbourhood quality, preferences for housing characteristics and subjective flood probabilities. Housing budgets are heterogeneous among agents. The capitalised discount for flood risk is calculated through analysis of sales during the simulations. Each time step, the realtor agent updates his price expectations through hedonic analysis and spatial interpolation (kriging) of sales in the previous time step.

12.3 Validation Techniques Used in Our Model

We can divide the validation techniques of our model into a theoretical part and an empirical part.

Theoretical:

- Economic theory: microeconomics, supply, demand and bidding, utility maximization

- Urban economics [15]: markets with spatial goods that are differentiated by distance to CBD, amenities and risks.
- Risk perception theory: risk negligence, prospect theory, expected utility

Empirical:

(a) For parameters on spatial and economic attributes:

- Two datasets in Eastern North Carolina – one coastal city with coastal flooding and one inland city with river flooding – with 4799 [11] and 3106 [10] property transactions, respectively
- Two datasets of actual housing properties and their structural characteristics in the same cities: 9793 properties in the Greenville dataset and 3481 properties in the Beaufort dataset.
- Empirical data on household income and budget spent on housing in the USA

(b) For parameters on behavioural attributes:

- Surveys among 600 current buyers and 600 current sellers

(c) For the structure of behavioural rules and interactions on the market:

- 2 × 2 hour in-depth interviews with real estate agents to specify the main architecture of the market (how ask and bid prices are formed, how agents negotiate prices, how they adjust prices, how learning on price expectations is happening) in 2013 and 2015
- 19 × half hour to 1-hour interviews with real estate agents in 2017

The model was initially validated against a set of stylised facts based on empirical observations on macroscale and interviews with two realtors to get an understanding on how the market should be modelled. Additionally we used microeconomic theories to engineer buyer and seller behaviour. After a few simulation studies, our demand for empirically informed behaviour on the microscale increased. We needed more data on the perceptions and considerations of individual households in order to better understand the underlying behaviour of agents in the market and to better link and understand it in the context of market responses to flood risk. Therefore we added the surveys (part b of empirical validation) at the last stage. These surveys allowed us to ask individuals specifically about their risk perception, their behavioural responses (e.g. would they actively avoid properties located in the flood zone?) and their willingness to pay for properties inside versus outside the flood zone.

12.4 Why Multiple Scales?

Our agent-based model was validated theoretically as well as empirically on multiple scales, from aggregated market-level responses to flooding events to individual responses to natural hazards. We applied a variety of approaches on multiple scales for a number of reasons. Firstly, we found that our theoretical models of risk perception in this agent-based land market model were insufficient to capture the price dynamics that were observed in empirical studies [17]. Secondly, empirical studies of market transactions that consist past flooding events were insufficient to extrapolate the behaviour of individual agents in a world where the climate is changing and severe flooding events become more common. Thirdly, we needed more data on the risk perceptions and choices made by individuals in order to understand the underlying behaviour of individuals that govern the observed market behaviour since neither of the stylised theoretical models can explain it well [ref to EE paper]. Hence, there seem to be additional factors that play a role in the actual decision-making.

The micro-macro validation approach allows for a better understanding of the system. ABMs allow us to study systemic behaviour by looking at how individual elements behave and interact. Any mismatch that we find between validated micro- and macro-level behaviour forces us to rethink the underlying mechanisms and relationships of the system's functioning. This iterative process is the core of the good modelling practice [18].

12.5 Contribution and Further Study

This paper contributes to our knowledge of validation in ABM in the following ways:

- It provides examples of ways to validate agent-based models: in particular, why it is useful to validate the model empirically on system level as well as on individual level.
- It provides some insights that help as a guide to other agent-based modellers.

We discussed our improvements in terms of empirical validation of micro-level validation. In the next stage of this research, the survey results will be analysed and implemented in the model. We assess the differences in outcomes when the model is based on validated system behaviour versus agent behaviour. Moreover, we will explore under which circumstance non-marginal shifts in property values may emerge when individuals experience a major flooding. It is of most interest to understand how social amplification of risk impacts this process and to what extent changes in probability and severity of floods matter.

Validation can be a continuous iterative process, and there is no definite answer as to how much empirical validation is enough in order to make a model useful for its purpose. Moreover, our experience in validating our model does not necessarily apply to other models as well, and the balance between theoretical and empirical validation may shift depending on the modelling purpose (e.g. predicting, explaining and understanding of real-world system dynamics). Most importantly, as with any model, the model design should make sense. It is up to expert judgement whether it does. It is not so much about how much the model itself needs to be validated but about how much information a person needs in order to be comfortable making well-reasoned modelling decisions.

References

1. Fagiolo, G., Moneta, A., Windrum, P.: A critical guide to empirical validation of agent-based models in economics: methodologies, procedures, and open problems. Comput. Econ. **30**(3), 195–226 (2007)
2. Windrum, P., Fagiolo, G., Moneta, A.: Empirical validation of agent-based models: alternatives and prospects. J. Artif. Soc. Soc. Simul. **10**(2), 8 (2007)
3. Anand, N., Meijer, D., Van Duin, J.H.R., Tavasszy, L., Meijer, S.: Validation of an agent based model using a participatory simulation gaming approach: the case of city logistics. Transp. Res. Part C: Emerg. Technol. **71**, 489–499 (2016)
4. Fagiolo, G., Guerini, M., Lamperti, F., Moneta, A., Roventini, A.: Validation of agent-based models in economics and finance. LEM Papers Series, 23 (2017)
5. Holm, S., Hilty, L.M., Lemm, R., Thees, O.: Empirical validation of an agent-based model of wood markets in Switzerland. PLoS One. **13**(1), e0190605 (2018)
6. Moss, S.: Alternative approaches to the empirical validation of agent-based models. J. Artif. Soc. Soc. Simul. **11**(1), 5 (2008)
7. Werker, C., Brenner, T.: Empirical calibration of simulation models. Papers on Economics and Evolution # 0410, Jena, Max Planck Institute for Research into Economic Systems (2004)
8. Abrahamse, W., Steg, L.: How do socio-demographic and psychological factors relate to households' direct and indirect energy use and savings? J. Econ. Psychol. **30**(5), 711–720 (2009)
9. Kahneman, D., Tversky, A.: Prospect theory: an analysis of decision under risk. Econometrica. 263–291 (1979)
10. Bin, O., Kruse, J.B., Landry, C.E.: Flood hazards, insurance rates, and amenities: evidence from the coastal housing market. J. Risk Insur. **75**(1), 63–82 (2008)
11. Bin, O., Landry, C.E.: Changes in implicit flood risk premiums: empirical evidence from the housing market. J. Environ. Econ. Manag. **65**, 361–376 (2013)
12. Quigley, J.M., Raphael, S.: Is housing unaffordable? Why isn't it more affordable? J. Econ. Perspect. **18**(1), 191–214 (2004)
13. Statista, http://www.statista.com/markets/411/topic/970/economy/
14. United States Census Bureau., http://www.census.gov/hhes/www/cpstables/032012/hhinc/toc.htm
15. Filatova, T.: Empirical agent-based land market: integrating adaptive economic behaviour in urban land-use models. Comput. Environ. Urban. Syst. **54**, 397–413 (2015)
16. De Koning, K., Filatova, T., Bin, O.: Improved methods for predicting property prices in Hazard prone dynamic markets. Environ. Resour. Econ. **69**(2), 247–263 (2018)
17. De Koning, K., Filatova, T., Bin, O.: Bridging the gap between revealed and stated preferences in flood-prone housing markets. Ecol. Econ. **136**, 1–13 (2017)
18. Jakeman, A.J., Letcher, R.A., Norton, J.P.: Ten iterative steps in development and evaluation of environmental models. Environ. Model Softw. **21**(5), 602–614 (2006)

Chapter 13
Towards Agent-Based Models of Rumours in Organizations: A Social Practice Theory Approach

Amir Ebrahimi Fard, Rijk Mercuur, Virginia Dignum, Catholijn M. Jonker, and Bartel van de Walle

Abstract Rumour is a collective emergent phenomenon with a potential for provoking a crisis. Modelling approaches have been deployed since five decades ago; however, the focus was mostly on epidemic behaviour of the rumours which does not take into account the differences between agents. We use social practice theory to model agent decision-making in organizational rumourmongering. Such an approach provides us with an opportunity to model rumourmongering agents with a layer of cognitive realism and study the impacts of various intervention strategies for prevention and control of rumours in organizations.

Keywords Rumour · Organization · Social practice theory · Agent-based model

13.1 Introduction

The phenomenon of rumourmongering has malicious impacts on societies. Rumours make people nervous, create stress, shake financial markets, and disrupt aid operations [27]. In organizations, rumours lead to unpleasant consequences such as breaking the workplace harmony, reduction of profit, drain of productivity, and damaging the reputation of a company [7, 20]. Recent work on the McDonald's wormburger rumour and the P&G Satan rumour confirm the negative impact of rumours on the productivity of firms [7].

For more than hundred years, scholars from a wide range of disciplines are trying to understand different dimensions of this phenomenon. Research in rumour studies can be classified according to the approach followed: a case-based approach and a model-based approach. In the case-based approach, results are based on case studies, not on models, making it hard to generalize their conclusions. The model-

A. E. Fard (✉) · R. Mercuur · V. Dignum · C. M. Jonker · B. van de Walle
Delft University of Technology, Delft, Netherlands
e-mail: A.EbrahimiFard@tudelft.nl; R.A.Mercuur@tudelft.nl; M.V.Dignum@tudelft.nl; C.M.Jonker@tudelft.nl; B.A.vandeWalle@tudelft.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_13

based approach tries to explain the phenomenon of rumours by model-based based simulations. The model-based approaches, so far, focus only on the dynamic of the spread, while rumour is a collective phenomenon, and the acts of individuals can influence the whole system. Rumours in organizations have been mainly approached with case-based studies and dynamic spreading model. To our knowledge there are no studies where the cognition of the individual is taken into account.

In our agent-based approach, we study the dynamics of the spread of rumours in organizations as an emergent (collective) behaviour resulting from the behaviour of individual agents using social practice theory. We use the proposed model to study the impact of change in organizational layout on control of organizational rumour.

The concept of social practices stems from sociology and aims to depict our 'doings and sayings' [24, p. 86], such as dining, commuting and rumourmongering. This paper uses the semantics of the social practice agent (SoPrA) model [19] to gain insights in rumourmongering in organizations.[1] SoPrA provides a unique tool to combine habitual behaviour, social intelligence and interconnected practices in one model. This makes SoPrA especially well-suited for studying the spread of rumours in organizations as this practice is largely habitual, social [10] and interconnected with practices as working and moving around. Given the lack of available empirical data on the social practice of rumourmongering, we give a proof of concept on how to collect data by doing eight semi-structured interviews.

This paper is organized as follows. The next section provides an overview of the research on rumours with an emphasis on studies of organizational rumour. Section 13.3 describes the context for our experiment and the methodology of data collection and data preprocessing. The model is introduced in Sect. 13.4. One possible experiment is described in Sects. 13.5, and 13.6 presents our conclusions, discussion and ideas for future work.

13.2 Background and Related Work

Rumours are unverified propositions or allegations which are not accompanied by corroborative evidence [7]. Rumours take different forms such as exaggerations, fabrications, explanations [23], wishes and fears [14]. Rumours have a lifecycle and change over the time. Allport and Postman in their seminal work psychology of rumour concluded that, as a rumour travels, it grows shorter, more concise, more easily grasped and told [12]. Buckner considers rumour a collective behaviour which is becoming more or less accurate while being passed on as they are subjected to the individuals' interpretations which depends on the structure of the situation in which

[1]Mercuur et al. [19] provides a static model of SoPrA based on literature and argued modelling choices. This paper applies this model to the domain and extends it by including competences and affordances and modelling a dynamic component based on [17]. Note that Mercuur et al. [19] is still under review and only available as pre-print at the moment of writing.

the rumour originates and spreads subsequently [4]. Rumours are conceived to be unpleasant phenomena that should be curtailed. Therefore, a number of strategies have been proposed to prevent and control them [4, 15, 22].

One of the rumour contexts that has received attention from researchers for almost four decades is organizations. Like rumour in general context which is explained in the above paragraph, rumour in organizations has different types and follows its own lifecycle [1–3, 7, 8, 13]. Also, to quell credible and non-credible organizational rumours, a number of different techniques and strategies have been suggested [7, 13]. The research approach also follows the same pattern, with a slight difference which to the best of our knowledge is qualitative without adopting any modelling approach.

The related literature reported above are based on case studies or experiments in the wild. This pertains to the types of rumour, dynamics of rumour and strategies to control rumours, either in general or in organizational contexts. These case studies and experiments are to inform the construction of theories and models underlying the phenomenon of rumourmongering. Theories and models, in turn, should be tested in case studies and simulations. Model-based approaches do just that. However, the current state of the art in model-based simulations of rumourmongering focuses only on the dynamics of the rumourmongering, comparable to the epidemic modelling and spread of viruses [6, 21, 26, 28, 31]. These models do not consider the complexities of the agents that participate in rumourmongering.

The research area of agent-based social simulations (ABSS) specializes on simulating the social phenomena as phenomena that emerge from the behaviour of individual agents. ABSS is a powerful tool for empirical research. It offers a natural environment for the study of connectionist phenomena in social science. This approach permits one to study how individual behaviour give rise to macroscopic phenomenon [9]. Such an approach is an ideal way to study the macro effects of various social practices, because it can capture routines which are practised by individuals on a regular basis in micro level and see their collective influence in a macro level.

13.3 Domain

This research investigates the daily routine of rumourmongering in a faculty building on the campus of a Dutch university. In this faculty, students, researchers and staff work in offices with capacity of one to ten people. Aside from the actual work going on in the building, filling a bottle with water, getting coffee from the coffee machine, having lunch at the canteen and going to the toilet are among the most obvious practices that every employee in this faculty does on a daily basis.

Nevertheless, there are other daily routines in the organization which are not that obvious. One of these latent routines is rumourmongering. Rumours or unverified information are transferred between students, researchers and staffs on a daily basis, during lunch, while queuing for coffee, when seeing each other in the hallways and

Table 13.1 Interviewees demography

Number of interviews	Number of different countries	Lowest educational level	Mean age	Female %
8	6	MSc	28	50

when meeting in classrooms and offices. All these situations are potential contexts for casual talks and information communication without solid evidence.

For data collection we conducted semi-structured explorative interviews with people from the above-mentioned faculty. Semi-structured interviews allow us to ask questions that are specifically aimed at acquiring the content needed for the SoPrA model, while still giving the freedom to ask follow-up questions on unclear answers. The data collection can be improved in future works by increasing the number of interviewees and diversifying them (not only asking from students). For demographic information, the reader is referred to Table 13.1. We prepared the following question set to ask from each interviewee based on the meta-model which will be explained in the next section:

1. What are the essential competencies for rumourmongering?
2. What are the associated values with rumourmongering?
3. What kind of physical setting is associated with rumourmongering?

Given the thin line between personality traits and competences, we used the Big Five model [11] to differentiate between personality traits and competences. For Question 2, we asked the interviewees to choose the relevant values from Schwartz's Basic Human Values model [25]. We asked the same set of questions about fact-based talk.

We processed the collected data in two ways before using it in the model. Firstly, we clustered answers that point to the same concept. For example, in Question 3, interviewees gave answers such as cafeteria, coffee shop and cafe to point to a place where people can get together and drink coffee. In the coffee example, we clustered answers under the term of "coffee place."

Secondly, we classified the answers to Question 2. As mentioned, for that question, we asked interviewees to pick associated values from Schwartz's Basic Human Values model. We used the third abstraction level of the model which is more fine-grained and compared to other levels and gave the interviewees a better idea of what they point to. However, a model based on level three would not allow us to compare the agents effectively. Therefore, we decided to wrap the answers and classify them based on second abstraction level. Using a classification based on the first abstraction level would have been too homogeneous in the sense that the agents would behave too similar, which would lose the effectiveness of the simulation.

13.4 Model

The model has two main parts: (i) static part and (ii) dynamic part. In the static part, the components of the model and their properties are described, and in the dynamic part, we explain the interaction of those components.

13.4.1 Static Part

This section describes the SoPrA meta-model which is used as the groundwork for our agent-based model, how we use empirical data to initiate the model, the model choices we make and how we tailor the model to the context of organization.

The SoPrA meta-model was introduced by Mercuur et al. [19] and describes how the macro concept of social practices can be connected to micro-level agent concepts. Figure 13.1 shows SoPrA in a UML diagram. The main objects in a SoPrA model are activities (e.g. fact talk, rumourmongering), agents (e.g. PhD students, supervisors), competences (e.g. networking, listening), context elements (e.g. office, cafeteria) and values. Values here refer to human values as found by the earlier stated Schwartz model, such as power or conformity. The social practice is an interconnection of (1) activities and (2) related associations as depicted by

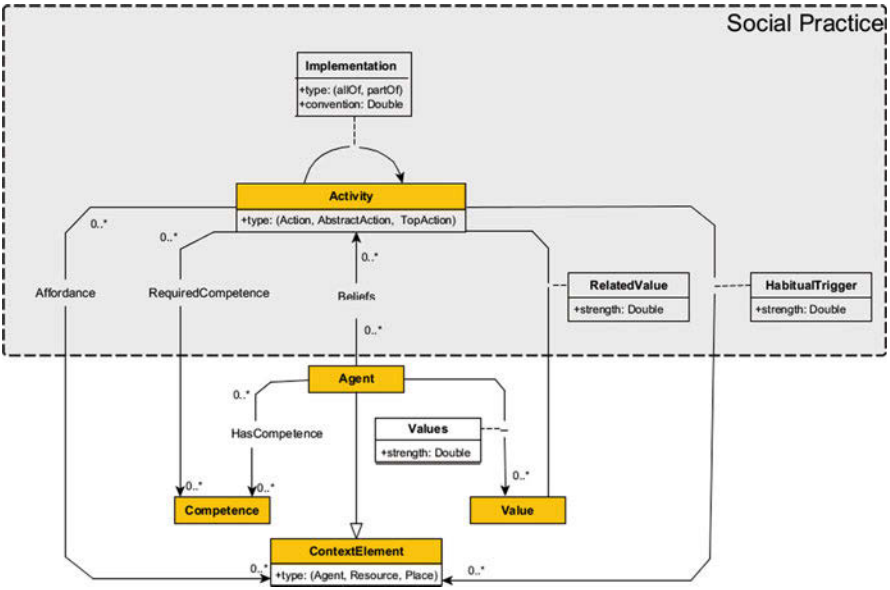


Fig. 13.1 The social practice meta-model captured in the Unified Modelling Language, including classes (yellow boxes), associations (lines), association classes (transparent boxes), navigability (arrow ends) and multiplicity (numbers)

Table 13.2 The associations attached to the activity and their specification

Association	Specification
Implementation	Which activities are a way of or a part of doing the activity
Affordance	Which context elements are needed to do the activity
RequiredCompetence	Which competences are needed to do the activity
Knowledge	Which activities an agent knows about
Belief	Which personal beliefs an agent has about the activity
RelatedValue	Which values are promoted or demoted by the activity
Trigger	Which context elements habitually start the activity
Convention	Which activities usually implement the activity

the grey box in Fig. 13.1. For example, the practice of talking consists of two possible activities: fact talk or rumourmongering. The social practice connects these different activities with the `Implementation` association. If activity A implements activity B, this means that A is a way of or a part of doing B.

The `Implementation` association is the first of several associations that are related to an activity (see Table 13.2). Most associations are fairly self-explanatory; however the `Trigger` and `Convention` attribute are a bit more complex. Following Wood and Neal [29], triggers are the basis for habitual behaviour. If an agent is near a context element that has a trigger association with an activity, then it will do that activity automatically (without, e.g. considering its values). Following Crawford and Ostrom [5], conventions are related to norms and signify that something is the normal way to do something. If an agent believes that activity A is a strategy for activity B, then it believes that other agents usually implement activity B by doing activity A.

The SoPrA meta-model does not only relate the activities to other classes, but the agent itself also has two types of associations, `HasCompetence` and `ValueAdherence` which play a role in choosing the activities it will do. The `HasCompetence` association links possible skills to the agent who masters those. The `ValueAdherence` association captures if an agent finds that value important.

The model can be initiated using empirical data. Note that in this study, we focussed on a small set of explorative interviews. We show with this initial data a proof of concept of how the model can be initiated. To properly ground the model, a larger and more rigorous empirical study is necessary.

The activity class has three instances: talking, rumourmongering and fact talk. The number of instances of agent can vary in the different experiments (see Sect. 13.5). The instances of the context element, competence and values class are based on the gathered data and can be found in Tables 13.3 and 13.4.[2] The

[2]The context element "Friend" and "Colleague" are special cases; these are rather attributes of context elements (i.e. agents) than context elements themselves. In our model these are to some extent implicitly captured, because the agents who one sees most often (i.e. friends, colleagues) are mostly likely to be habitually associated with an action.

Table 13.3 The elements associated with the rumourmongering activity

Rumourmongering		
Context elements	Meaning	Competence
Friend	Self-direction	Sneaky skills
Coffee place	Power	Network skills
Hallway	Hedonism	Talking skills
Restaurant	Achievement	Observing skills
Office	Benevolence	
Phone		
Computer		

Table 13.4 The elements associated with the fact talk activity

Fact talk		
Context elements	Meaning	Competence
Colleague	Universalism	Being knowledgeable
Academic staff	Self-direction	Listening skills
Office	Benevolence	Critical thinking skills
Conference	Achievement	Communication skills
Meeting room	Tradition	
Classroom		
Restaurant		
Phone		
Computer		
Pen		
Coffee		

complete static model consists both of object instances and associations between these instances. An example focusing on one agent (i.e. `Bob`) and one activity (i.e. `rumourmongering`) is shown in Fig. 13.2. Bob believes that the activity of rumourmongering is related to the value of privacy, curiosity and social power. He thinks it requires the competence of networking and noticing juicy details and thinks the activity is triggered (to some extent) by the hallway, restaurant and another agent named Alice. Furthermore, he himself has the competence of networking and adheres strongest to the value of ambition and weakest to the value of pleasure.

The agents differ in which activity they associate with which element. In other words, the SoPrA meta-model does not initiate one social practice that all agents share, but one social practice *per agent*. The chance that an agent relates an activity to a competence is based on the empirical data we gathered in the interviews. For example, if 50% of the interviewees linked critical thinking skills to fact talk, the chance an agent makes this association depends on a binomial distribution with $p = 0.5$. For `relatedValue` association and `HabitualTrigger` association, all agents make the associations as mentioned in Tables 13.3 and 13.4. However, the weights differ per agent. The weights for the `relatedValue` association are picked from a normal distribution between 0 and 1. Given the lack of empirical data on the relation between activities and human values, we follow the related

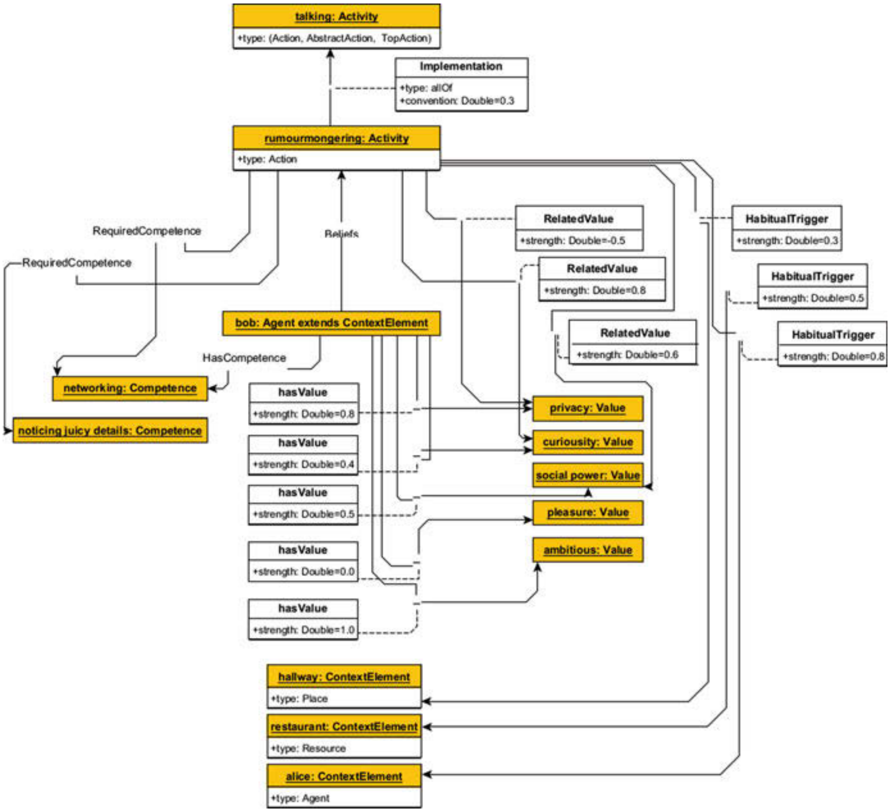


Fig. 13.2 An instance of the SoPrA meta-model for the activity of rumourmongering and one agent. For illustration purposes, the associations related to the activity talking and the agent "Alice" are omitted

finding of the World Values Survey that people adhere to values with roughly a normal distribution [30]. The weights for `HabitualTrigger` are picked on a logarithmic distribution based on the empirical work of [16]. One interesting modelling choice we made was to drop the `Affordance` associations in the conceptual model. The SoPrA meta-model conceptualizes two associations with context elements. The `HabitualTrigger` association representing that some context element can automatically lead to a reactive action and the `Affordance` association representing that some context elements are a pre-condition to enact a certain behaviour. None of our interviewees mentioned a possible context element that affords rumourmongering fact talk. As such this association seemed irrelevant for our model.

The associations related to the agents themselves are based on random distributions. Each competence has a 50% chance to be related to an agent. Each value is associated to each agent, but the weights differ. The weights for the `hasValue` association strength are based on a correlated normal distribution. Schwartz [25] shows that the strength to which people adhere to values is correlated. For example, people who positively value universalism usually negatively value achievement. We use the correlations found by Schwartz [25] to simulate intercorrelated normal distribution from which we pick the weights. In future work, we aim to extend our interviews to also gather data that can inform these weights.

For our modelling context, we need to extend the SoPrA model with a spatial component. We do this by adding two attributes to the `ContextElement` class called `x-coordinate` and `y-coordinate`. These coordinates can be used by the agent to sense which objects are near. Note that every agent is also a context element as indicated with the "generalization" association in the UML diagram.

13.4.2 Dynamic Part

This section describes the dynamic part of the model which on each tick comprises:

1. An agent decides on its location using the moving submodel and updates its coordinate attributes.
2. An agent decides if it will engage in fact talk or rumourmongering based on the choose-activity submodel.

The moving submodel has four components that agents can transfer between. As it is shown in Fig. 13.3, the initial state is offices, and from that state agents can leave their offices and pass the hallway to either have lunch at the restaurant or grab a cup of coffee at the coffee place. During the interviews, we discovered most of the people do those daily routines around the same period of time, and only a few people do not follow this pattern and leave their offices out of usual time periods, so we concluded the transition of agents between different locations is a random phenomenon which follows a normal probability distribution.

The choose-activity submodel is based on Mercuur et al. [18] and has three stages. The submodel is depicted in Fig. 13.4. The agent starts by considering both

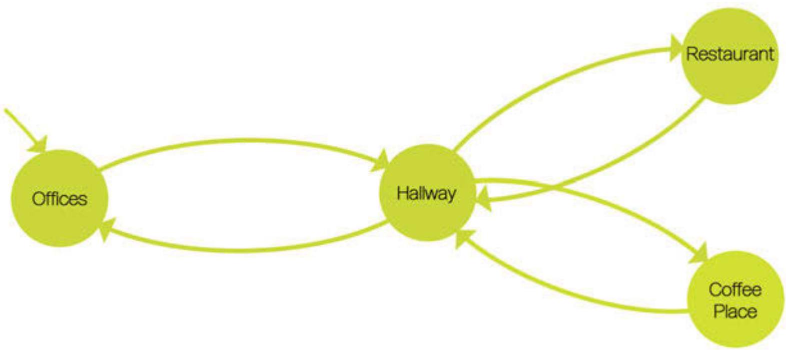


Fig. 13.3 The moving model for agents

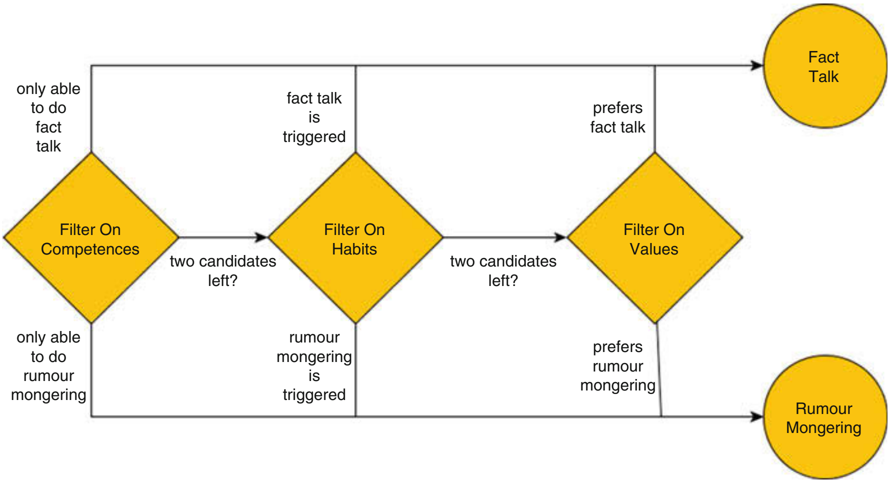


Fig. 13.4 The choose-activity submodel and the three stages the agent uses to decide on its activity: competences, habits and values

rumourmongering and fact talk. At each stage the agent makes a decision on one cognitive aspect. If this aspect is not conclusive, it will prolong the decision to the next stage. In the first stage, the agent compares its own competences to the competences that it believes to be required for the activity. In our example model depicted in Fig. 13.2, Bob would decide it cannot do the activity of `rumourmongering`, because it requires a competence he does not have: noticing juicy details. As such, Bob will engage in fact talk. (Note that if Bob does not have the skill to do either activity, then the decision is also prolonged to the next stage.) In the second stage, an agent tries to make a decision based on its habits. It will survey its context and decide which context elements are near, i.e. resources, places or other agents. If it has a habitual trigger association with a particular strong strength between one of those context elements and either rumourmongering or fact talk, it will automatically do that action. In the last stage, the agent will consider how strongly it relates certain values to both activities and how strongly it adheres itself to these values. Consequently, it makes a comparison between the two activities and decides which best suits its values. For the complete implementation of the habitual model and value model, we refer to [17].

13.5 Experiment

The proposed rumour model with elements associated with physical settings, individual values and competencies enables us to investigate impacts of a variation of settings and interventions on the spread of rumours in organizations.

One of the open questions in organizational rumour literature is the effectiveness of different prevention and control strategies. In our approach we only need to extend the model with the specific elements and characteristics of the case that we would like to study. In this paper we study the effect of organizational layout on rumour dynamics. In our case, we take the size of offices and number of coffee places as the proxies for organizational layout and juxtapose two organizational layouts cases (Fig. 13.5) to understand the impact of layout on rumourmongering dynamic.

To set up the model, we determine the number of agents and then initialize the context and agents. In the organization that we studied, each section has on average 50 people; therefore, we pick 50 as the number of the agents. For context initialization, we design the layouts and assign agents to different locations, and then we initialize agents with probability distributions for routines such as grab a cup of coffee or having lunch. After the model setup, it can be executed.

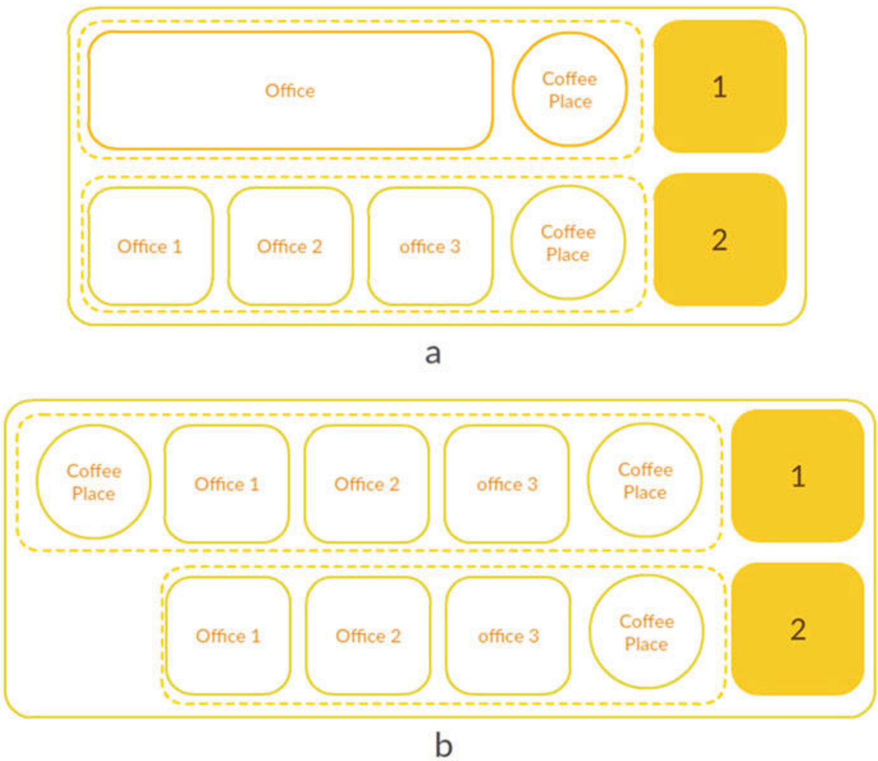


Fig. 13.5 (**a**) In this case, we study the impact of office size on dynamics of rumourmongering. (**b**) In this case we study the impact of number of coffee places on the dynamics of rumourmongering

13.6 Discussion and Future Research

Modelling rumourmongering has been studied since 1964. So far, the modelling did not consider the complexities of individual agents and mostly focused on the spreading behaviour of the phenomenon. In the model proposed in this paper, agents have a cognitive layer that deploys social practice theory and views rumour as a routine with associated competencies, values and a physical setting.

In this research, we narrowed our study to the context of organization, and after introducing the generic model, we tailored our model to the context of organization via empirical data collected through interviews conducted in a Dutch university. Based on explorative interviews, we established that social practice theory is likely to be applicable as people shared a view on rumour, and their habits regarding rumour and rumours seem to be intertwined with other activities.

Our model can be used to study a wide range of topics in organizational rumour studies, in particular for testing the effectiveness of interventions for prevention and control of rumours in organizations.

Future work is to extend the questionnaire by asking about associations, conduct more and more rigorous interviews, implement the model and run the proposed experiments that explore different organization layouts. Furthermore, we aim to validate our model by looking at how rumours travel from person to person in the organization during a pre-selected time period.

Contributions and Acknowledgement Ebrahimi Fard and Mercuur wrote the first draft. Ebrahimi Fard provided the domain knowledge and collected most data, whereas Mercuur provided the meta-model and methodological knowledge. Dignum, Jonker and van der Walle supervised the process and contributed to the draft by providing comments, feedback and rewriting. This research was supported by the Engineering Social Technologies for a Responsible Digital Future project at TU Delft and ETH Zurich.

References

1. Bordia, P., Hobman, E., Jones, E., Gallois, C., Callan, V.J.: Uncertainty during organizational change: types, consequences, and management strategies. J. Bus. Psychol. **18**(4), 507–532 (2004)
2. Bordia, P., Jones, E., Gallois, C., Callan, V.J., Difonzo, N.: Management are aliens! rumors and stress during organizational change. Group Org. Manag. **31**(5), 601–621 (2006)
3. Bordia, P., Kiazad, K., Restubog, S.L.D., DiFonzo, N., Stenson, N., Tang, R.L.: Rumor as revenge in the workplace. Group Org. Manag. **39**(4), 363–388 (2014)
4. Buckner, H.T.: A theory of rumor transmission. Public Opin. Q. **29**(1), 54–70 (1965)
5. Crawford, S.E.S., Ostrom, E.: A grammar of institutions. Polit. Sci. **89**(3), 582–600 (2007). http://www.jstor.org/stable/2082975
6. Daley, D.J., Kendall, D.G.: Epidemics and rumours. Nature **204**(4963), 1118 (1964)
7. DiFonzo, N., Bordia, P., Rosnow, R.L.: Reining in rumours. Organ. Dyn. **2**, 47–62 (1994)
8. DiFonzo, N., Bordia, P.: A tale of two corporations: managing uncertainty during organizational change. Hum. Resour. Manag. **37**(3–4), 295–303 (1998)

9. Epstein, J.M.: Agent-based computational models and generative social science. Gener. Soc. Sci. Stud. Agent Based Comput. Model. **4**(5), 4–46 (1999)
10. Gersick, C.J., Hackman, J.R.: Habitual routines in task-performing groups. Organ. Behav. Hum. Decis. Process. **47**(1), 65–97 (1990)
11. Goldberg, L.R.: The structure of phenotypic personality traits. Am. Psychol. **48**(1), 26–34 (1993)
12. Hart, B.: The psychology of rumor. Psychiatry Interpersonal Biol. Process. **28**, 1–26 (1916)
13. Kimmel, A.J.: Rumors and Rumor Control: A Manager's Guide to Understanding and Combatting Rumors. Lawrence Erlbaum, Mahwah (2004)
14. Knapp, R.H.: A pyschology of rumour. Public Opin. Q. **8**(1), 22–37 (1944)
15. Knopf, T.A.: Beating the rumors: an evaluation of rumor control centers. Policy Anal. **1**(4), 599–612 (1975)
16. Lally, P., Van Jaarsveld, C.H.M., Potts, H.W.W., Wardle, J.: How are habits formed: modelling habit formation in the real world. Eur. J. Soc. Psychol. **40**(6), 998–1009 (2010)
17. Mercuur, R.: Interventions on contextualized decision making: an agent-based simulation study. Ph.D. thesis, Utrecht University (2015). https://dspace.library.uu.nl/handle/1874/323482
18. Mercuur, R., Dignum, F., Kashima, Y.: Changing habits using contextualized decision making. In: Advances in Intelligent Systems and Computing, vol. 528 (2017)
19. Mercuur, R., Dignum, V., Jonker, C.M.: Modelling agents endowed with social practices: static aspects. ArXiv e-prints, pp. 1–22 (2018). http://arxiv.org/abs/1811.10981
20. Michelson, G., Mouly, S.: Rumour and gossip in organisations: a conceptual study. Manag. Decis. **38**(5), 339–346 (2000)
21. Nekovee, M., Moreno, Y., Bianconi, G., Marsili, M.: Theory of rumour spreading in complex social networks. Phys. A Stat. Mech. Appl. **374**(1), 457–470 (2007)
22. Oh, O., Agrawal, M., Rao, H.R.: Community intelligence and social media services: a rumor theoretic analysis of tweets during social crises. MIS Q. **37**(2), 407–426 (2013)
23. Prasad, J.: The psychology of rumour: a study relating to the great Indian earthquake of 1934. Br. J. Psychol. Gen. Sect. **26**(1), 1–15 (1935)
24. Schatzki, T.R.: Social Practices. A Wittgensteinian Approach to Human Activity and the Social. Cambridge University Press, Cambridge (1996)
25. Schwartz, S.H.: An overview of the Schwartz theory of basic values. Online Read. Psychol. Cult. **2**, 1–20 (2012)
26. Turenne, N.: The rumour spectrum. PLoS One **13**(1), e0189080 (2018)
27. Vosoughi, S., Roy, D., Aral, S.: The spread of true and false news online. Science **359**(6380), 1146–1151 (2018)
28. Wang, C., Tan, Z.X., Ye, Y., Wang, L., Cheong, K.H., Xie, N.G.: A rumor spreading model based on information entropy. Sci. Rep. **7**(1), 9615 (2017)
29. Wood, W., Neal, D.T.: A new look at habits and the habit-goal interface. Psychol. Rev. **114**(4), 843–863 (2007)
30. World Values Survey Association: World Values Survey. http://www.worldvaluessurvey.org/wvs.jsp
31. Zanette, D.H.: Dynamics of rumor propagation on small-world networks. Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics **65**(4), 9 (2002)

Chapter 14
Fixing Sample Biases in Experimental Data Using Agent-Based Modelling

Mike Farjam and Giangiacomo Bravo

Abstract We present how agent-based models can be used to correct for biases in a sample. The approach is generally useful for behavioural experiments where participants interact over time. The model we developed copied mechanics of a behavioural experiment conducted earlier, and agents in the model faced the same strategic choices as human participants did. We used the data from the experiment to calibrate agent behaviour such that agents reproduced patterns observed in the experiment. After this learning phase, we resampled agents such that their characteristics (political orientation) were similar to those found in the real world. We found that after the correction for the bias, agents produced patterns closer to those commonly found.

Keywords Agent-based modelling · Experiment · Bias · Methodology

14.1 Introduction

The goal of this work is to understand whether it is possible to use agent-based modelling to correct for biased samples in behavioural experiments. As interaction typically occurs over multiple rounds and participants receive feedback on and react to the behaviour of others, each individual does not represent an independent unit, which makes it impossible to correct for the bias by taking a stratified resample of the data originally collected. Different group compositions may indeed lead to very different dynamics, which cannot be captured by the resampling. We propose here to correct sample biases using an ABM where the agents' parameters are based on the data from the experiment, with the resampling that occurs at the agent level while the group dynamics is simulated by the model. By using data from a collective-risk

M. Farjam (✉) · G. Bravo
Centre for Data Intensive Sciences and Applications, Linnaeus University, Växjö, Sweden
e-mail: Mike.Farjam@lnu.se

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_14

social dilemma (CRSD) experiment [1, 2], we show that a resampled ABM using more representative group compositions produces results that are closer to the ones from established research in the field than the outcome of the biased experiment.

14.2 The CRSD Experiment

Milinski et al.'s [3] CRSD has become the leading protocol to experimentally study climate change mitigation behaviour. Played in groups of six participants over ten rounds, in each round, participants decide whether to keep (part of) their endowments or to contribute to a "climate protection" account. At the end of the round, they receive feedback about the other group member's contributions. Money on the climate protection account is used to pay for pro-climate actions in the real world. If the total contribution of the group is above a fixed threshold, the money remaining in the private account is paid to each participant. If the total contribution is lower, all the private account money is lost with a certain probability.

More precisely, each participant is first endowed with 40 experimental currency units (ECU) and can contribute {0, 2, 4} ECU to the climate account in each round. The probability of losing the private account money if a threshold of 120 ECU in the climate protection account is not reached is referred to as p and is $\{0.1, 0.5, 0.9\}$. Milinski et al. showed that, using a sample of German students in a university lab, all groups in the $p = 0.1$ condition, 90% of the groups in the $p = 0.5$ one and 50% of the groups in the $p = 0.9$ one *failed* to successfully solve the dilemma. Similar findings were produced by subsequent lab experiments using similar participants' populations [2, 4].

Being interested in checking whether these results were robust to a change of the participants' population, we recently replicated Milinski's experiment using a sample of adult US citizens recruited through mTurk, a common recruitment tool for behavioural experiments [5]. To our surprise, we recorded much higher contribution levels, leading to only 19%, 10% and 7% of the groups failing to reach the threshold in the $p = 0.1$, $p = 0.5$ and $p = 0.9$ conditions, respectively. Deeper analyses showed that, while the age and gender distributions of the participants were more or less comparable to the US ones, our sample over-represented left-wing, environmental-concerned people. Both variables were measured on a 0–10 point scale in post-experimental questionnaires and showed medians of 4 (left wing 0–right wing 10) and 8 (low concern 0–high concern 10), respectively. The median values for similar questions in the US sample of the World Value Survey (WVS), wave 2010–2014, were 5.5 for both political orientation and environmental concern[1] (Fig. 14.1).

The importance of these variables is confirmed by the fact that our measure of environmental concern and, to a lower extend, of left-wing orientation significantly correlated with contributions ($r = 0.44$ and $r = 0.17$, respectively). In addition,

[1]To help the comparison, the original values were rescaled into our 11-point scale.

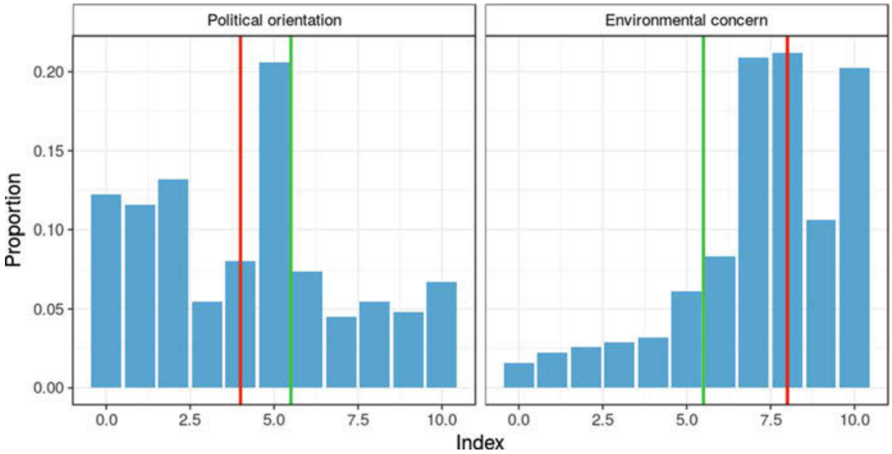


Fig. 14.1 Distribution of the political orientation and environmental concern scales among participants in the mTurk experiment. Median values in the experiment are indicated as red lines, actual USA values from the WVS as red lines

knowing that the contributed money was used to buy carbon compensation credits, some participants may have taken this opportunity to produce positive real-world externalities.[2] Also taking into account the strong political polarization of the topic in the USA, this probably led to a much higher willingness to contribute than in standard laboratory experiments.

14.3 Methods

We developed an ABM in NetLogo 6.0 [6] reproducing the CRSD game dynamics. Agents' data mapping the experiment participants are first uploaded in the model, and then the game goes on as in the original experiments. In each period, agents choose a level of contribution following the procedure described below. Then contributions are aggregated for each group, and agents are "informed" of the total group contribution and the remaining distance from the threshold. After 10 periods, corresponding to the 10 experimental rounds, the proportion of groups that reached the threshold is measured.

To determine the behaviour of agents starting from the empirical data, we estimated a linear mixed effects model. The political orientation and environmental concern scales were used to predict the contribution in the experiment in each period, along with the value of p used for that specific group, a measure of inequality in contributions (the ratio between the subject contribution and the mean

[2]Following the protocol, 800$ were used to offset the emission of 2050 t CO_2.

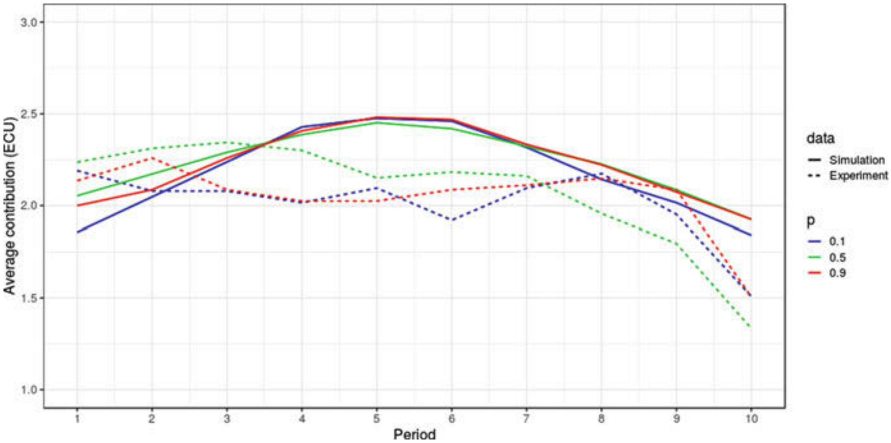


Fig. 14.2 Contribution dynamics in the empirical and simulated data

contribution of other group members in the previous period) and the distance from the threshold weighted by the period number. The resulting coefficients, along with the subject characteristics, such as political orientation and environmental concern, were then used to predict individual contributions. To transform the continuous output of the linear model into the discrete set of possible contributions, two cutting points were estimated by calibrating the model in order to maximize its capacity to predict the success of each group.

The resulting model was able to reproduce well the outcome of the experiment: predictions were correct for 70/70 groups that reached the threshold and 8/9 groups that did not. In addition, the model was able to capture the declining dynamics of contributions in the final part of the experiment, once most group passed the threshold, although it slightly over-estimated contributions in the central part of the game (Fig. 14.2).

To resample the agent population, we implemented a simple NetLogo procedure where one agent with a value of political orientation below the actual US median was randomly deleted and replaced by a copy of another agent having political orientation above the median. The procedure was repeated until the median in the simulation reached the actual US level. The same occurred for the environmental concern variable, except that agents with high concern were deleted and replaced by ones with low concern. Once completed the resampling procedure, the CRDS game was played as above, except that all groups were reshuffled. One hundred repetitions of the model were run for each level of p and the corresponding proportion of groups reaching the threshold recorded. On average, 38.4% of the groups in $p = 0.1$ failed to reach the threshold, 35.3% in $p = 0.5$ and 31.7% in $p = 0.9$, with a clear decrease of the success cases in comparison with the mTurk experiment results.

14.4 Conclusions

The outcome of our resampling model were much closer to the ones in Milinski's experiment [3], at least in the $p = 0.9$ condition. However, the same did not fully occur in the other two conditions condition where, although the model led to a significant decrease in comparison with the mTurk experiment, it still predicted that around 60% of the groups would pass the threshold, while none did so in Milinski's $p = 0.1$ condition and only 10% in the $p = 0.5$ one [3].

One reason for this could be that our model is fully data-driven and, as a consequence, has limits in correctly predicting behaviour that is too far from the one observed in the original sample. From this point of view, a model based on sounder cognitive mechanisms could improve the result. On the other hand, students often exhibit more selfish behaviour in experiments than older participants [7], which means that our model outcome could fit the actual behaviour of the general US population better than Milinski's one based on a sample of German students. In addition, subsequent experiments that reproduced or slightly modified Mikinski's design recorded higher contributions than the original study [4, 8]. For instance, Tavoni et al. [8], who replicated Milinski's $p = 0.5$ treatment, had 50% of the groups failing to reach the threshold: a clear decrease in comparison with Milinski and a result much closer to our resampled-group outcome.

While it is difficult to solve this issue without running another experiment on an appropriate sample, the method we illustrated in this paper could be more carefully tested by splitting existing experimental dataset into different samples and trying to use one part to predict the other. Future work will focus on doing this.

References

1. Farjam, M., Nikolaychuk, O., Bravo, G.: Does risk communication really decrease cooperation in climate change mitigation? Clim. Chang. **149**(2), 147–158 (2018)
2. Farjam, M., Nikolaychuk, O., Bravo, G.: Experimental evidence of an environmental attitude-behavior gap in high-cost situations. Ecol. Econ. **166**, 106–434 (2019)
3. Milinski, M., Sommerfeld, R.D., Krambeck, H.J., Reed, F.A., Marotzke, J.: The collective-risk social dilemma and the prevention of simulated dangerous climate change. Proc. Natl. Acad. Sci. **105**(7), 2291–2294 (2008)
4. Dannenberg, A., Löschel, A., Paolacci, G., Reif, C., Tavoni, A.: On the provision of public goods with probabilistic and ambiguous thresholds. Environ. Resour. Econ. **61**(3), 365–383 (2015)
5. Paolacci, G., Chandler, J., Ipeirotis, P.G.: Running experiments on Amazon Mechanical Turk. Judgm. Decis. Mak. **5**(5), 411–419 (2010)
6. Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling. Northwestern University, Evanston (1999)
7. Belot, M., Duch, R., Miller, L.: A comprehensive comparison of students and non-students in classic experimental games. J. Econ. Behav. Organ. **113**, 26–33 (2015)
8. Tavoni, A., Dannenberg, A., Kallis, G., Löschel, A.: Inequality, communication, and the avoidance of disastrous climate change in a public goods game. Proc. Natl. Acad. Sci. **108**(29), 11825–11829 (2011)

Chapter 15
Simulation of Behavioural Dynamics Within Urban Gardening Communities

Arthur Feinberg, Elena Hooijschuur, and Amineh Ghorbani

Abstract In order to better understand the mechanisms leading to resilient urban gardening systems, we revisit Ostrom's institutional design principles with an agent-based model (ABM) and implement behavioural dynamics as structured by the theory of reasoned action. Our experiments show that sanctioning bad behaviour in general increases the group cohesion and leads to longer collective action. Higher success rates occur for cases in which volunteers join for socialising rather than just taking crop yield. However, the design principles are not blueprints leading de facto to robust gardening systems: the combination of these principles is instead determining.

Keywords Design principles · Individual behaviour · Urban gardening

15.1 Introduction

Urban citizens engage more and more in neighbourhood-scale collective action such as urban gardening. Unfortunately, many of these initiatives fail in the first 5 years either for internal or external reasons. In this work we focus on the internal institutions and behavioural dynamics of individual participants leading to a more prosperous initiative. As in all common-pool resources (CPRs), it is difficult or undesirable to exclude people from the garden, and the consumption of one user diminishes the possibilities of other users [8]. Sometimes volunteers take a share of the resource at no cost: this is called "free-riding" [6]. This issue can diminish the willingness of the whole group to contribute, which harms the functioning of collective management [1].

A. Feinberg (✉) · E. Hooijschuur · A. Ghorbani
Delft University of Technology, Delft, The Netherlands
e-mail: a.feinberg@tudelft.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_15

Through many case studies of collective CPR management, E. Ostrom found eight design principles (DPs) leading to robust collective action. The DPs are not blueprints that are applied to all situations in the same way. Rather, they are structural similarities that are found in self-organised systems whose members have adapted and learned to be robust to social, economic and ecological disturbances [8, 9]. An existing study has investigated the effects of the DPs on collective action in urban gardens in the UK [1]; another study investigated the dynamics of collective action with the use of a system-dynamics model, without DPs [2]. What exact role Ostrom's DPs have on the robustness of systems is not clear yet. We intend to elucidate this point with the help of social simulations by focusing on the system of a community garden.

15.2 Methods

We gather theoretical information from literature on community gardens, robust collective action, behaviour dynamics and agent-based modelling (ABM). Our agents are people volunteering to care for urban community gardens. Practical information is collected on an urban gardening case in Rotterdam (NL) and via a database of over 120 community gardens in Germany [10]. Both the case study and the database, along with Chalise's work [2], provide insights on the volunteers' motivations to go gardening. In addition we extract from the database various community characteristics (e.g. group size) and types of interactions between volunteers. We learn from the case study how to implement Ostrom's DPs in our model. Only individual decision-making is relevant in our study.

Our ABM methodology is inspired by the procedure of van Dam et al. [11]. We build our model with the overall structure of the Institutional Analysis and Design (IAD) framework [8]: external variables (*biophysical conditions*, *attributes of community* and *rules-in-use*) determine the *action situations* taken by the members of a system; the resulting *interactions* and their *outcomes* are evaluated to update the external variables and the actions taken. In our case the *biophysical conditions* and *attributes of the community* boxes are defined thanks to structured interviews in our case study and to the database [10]. The *rules-in-use* box reflects in our case which of the Ostrom's DPs are manifesting in the system and how. For each agent, taking action is evaluated with the formalisation of behaviour dynamics defined in the Theory of Reasoned Action [3, 5], shown in Fig. 15.1: a resulting behaviour depends both on attitudes and subjective norms. A normative belief's strength reflects action reciprocity [7]. In our system the agent's behaviour (*action situations* box) is affected by the *rules-in-use*, subjective norms, the *evaluation* of previous *interactions* and their related *outcomes*.

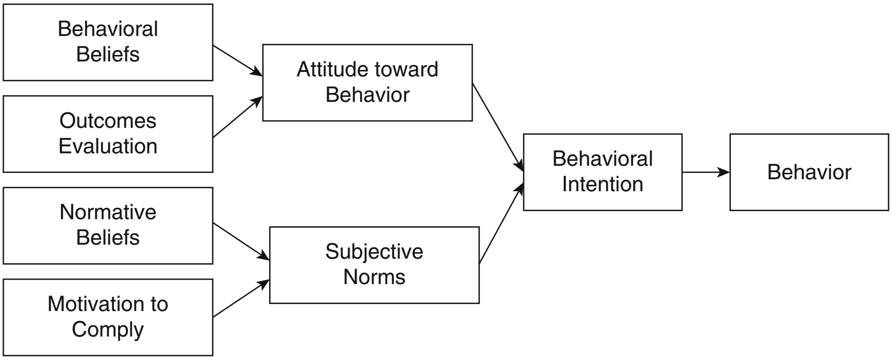


Fig. 15.1 Overview of the theory of reasoned action [5]

15.2.1 Concept Formalisation

We implement practical variables (see Fig. 15.2) from Ostrom's DPs which are both operational and relevant to community gardening. In this sense we have ruled out the following principles: minimal recognition of assembly and the nestling of enterprises. We define the three possible action situations: volunteering, taking garden yield and violating a rule. We isolate from literature the main motivations for volunteering, label them and give a related *belief strength* value and a *belief evaluation* value to each agent. The first is fixed while the second is updated at every round. The individual motivations for gardening are social cohesion and development, enhancing cultural practices, garden yield (to consume fresh food or save money), enjoying gardening (to be outdoors or for spiritual practices), sustainability, education, urban land accessibility and improving health. Negative motivations are the uncomfortable conditions (e.g. bad weather) or the excessive amount of work required. These motivations affect the agent's attitude, based on past personal experiences. From the group standpoint, the perceived need for contribution to the garden constitutes the group normative belief. The motivation to comply with it generates the subjective norm. In a given round, the probability for volunteering depends on the weighted sum of the attitude and subjective norm tested against a *contributing threshold*. The weights come from a study which evaluates the participation to team sports [4]. Taking yield and violating a rule are regulated by a set probability. The rules controlling these actions directly depend on the application of the DPs variables. The concept model is then translated into computer language using Netlogo.

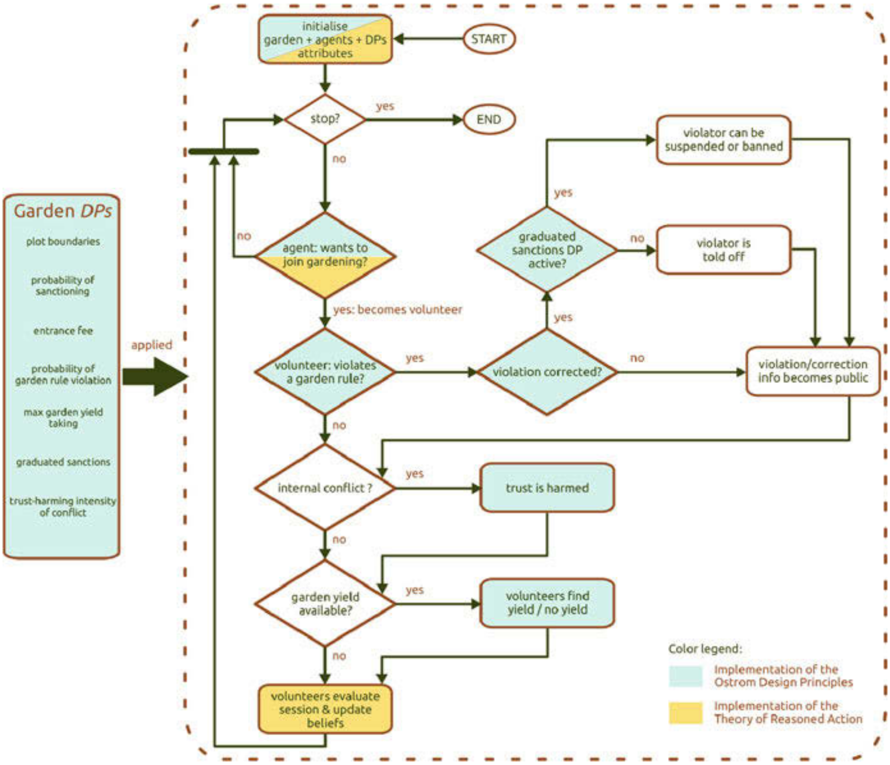


Fig. 15.2 Narrative model of the urban gardening case with applied design principles

15.2.2 Model Simulation

Due to our large number of parameters, we proceed to Latin hypercube sampling [11] to define the most relevant parameter spaces. Each experiment is run 100 times for repetition. This amount of repetitions is implemented in the NetLogo behaviourSpace by varying a randomSeed, so that each experiment can be reproduced.

15.2.3 Outputs Analysis and Validation

We measure the robustness of a scenario by the number of simulation ticks required until only one or no volunteer is left in the garden. We also measure the overall trust among volunteers and their beliefs for social cohesion and for yield taking. The simulation outputs are analysed with correlation tables and decision trees from

R package *ctree*. We validate the model in two ways: (1) a historic replay with the case study and (2) an expert validation. The expert is the author of the database mentioned above.

15.3 Findings

We found globally that higher probabilities of sanctioning increase the robustness of our system and the trust among volunteers. In addition, the higher the probability for a rule to be violated, the lower the trust. In terms of motivation, the system performed longer when volunteers did so for social cohesion instead of simply taking yield from the garden. This is also confirmed by the fact that higher cohesion values are found when sanctioning is in place. Taking more than its share of yield badly impacts the overall belief for yield taking, while the presence of plot boundaries around the garden increases this belief. However the use of decision trees showed us, for example, that with another combination of DPs, high probability of sanctioning may reduce robustness. This suggests that it is not DPs alone but a combination of them that determines the robustness of the collective action. We also observed that the system is more vulnerable to volunteers taking more than their share when poor conflict resolution mechanisms are in place. Figure 15.3 points to this sensitivity: collective action is more likely to collapse with higher trust-harming conflicts (blue line on the left in the tree). In our historic replay, the collective action within the case study indeed collapsed for this very reason: a

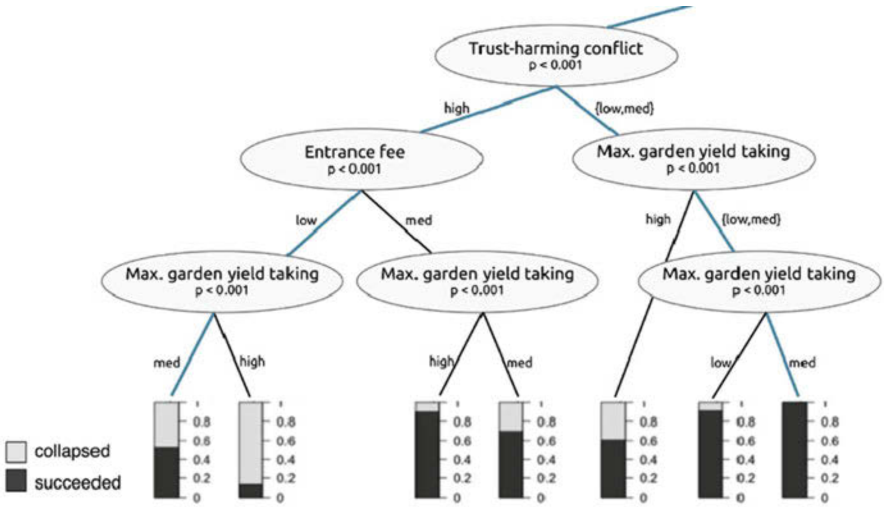


Fig. 15.3 Part of a decision tree: the left blue line represents the Rotterdam case

conflict emerged, which could not be dealt with efficiently and eventually escalated causing the amount of volunteers to dramatically decrease.

All the results are confirmed by our expert except the link between effective sanctioning and higher trust: this is because in our model, we did not include the possibility for people to volunteer solely with benevolent intentions. Instead we assumed that all volunteers make mistakes or violate rules with a given probability which evolves over the simulations in a way to mimic the others' behaviour. Also, our expert doubts the negative effect of taking too much yield on the belief for yield taking, probably because in our expert's samples, yield was not a strong motivation for joining in gardening activities.

15.4 Conclusions

Despite modelling assumptions and simplifcations, the ABM developed in this work contributes to the ongoing research on individual decision-making mechanisms. It also provides insights on making agents more realistic without excessive personalisation nor contextualisation. Secondly, this research adds on to the discussions around Ostrom's Design Principles. We found that collective action in urban gardens is overall facilitated by monitoring plot boundaries and group boundaries. However, Ostrom's Design Principles should not be seen as standalone miracle ingredients, as she warned herself at a dedicated workshop [9]: we have shown in our simulations the importance of Design Principles combinations, leading to a successful or failed initiative. This result had previously been found empirically [1]. For this purpose, decision trees are of great help in analysing the optimal and context-specific combination of institutions at play. The decision trees have proved to be useful and intuitive to communicate the results to the local actors. They can be used to reinforce collaborations with the municipality. We therefore highly recommend for social modellers to present intuitive graphical results to communicate with their objects of study, especially if they can help the modeller refine its model through historic replay.

References

1. Butler, A.: Collective action in UK urban community growing projects: a comparative analysis. Master's thesis (2013)
2. Chalise, N.: Collective Action Dynamics in Urban Neighborhoods: A Study of Urban Community Gardens. Washington University, St. Louis (2015)
3. Darnton, A.: Behaviour change knowledge review–reference report. Andrew Darnton for the Government Social Research Unit (2008)
4. Eves, F., Hoppéa, R., McLaren, L.: Prediction of specific types of physical activity using the theory of planned behavior. J. Appl. Biobehav. Res. **8**, 77–95 (2007)

5. Fishbein, M., Ajzen, I.: Predicting and changing behavior: the reasoned action approach. Taylor & Francis, New York (2011)
6. Moreira, J.A., Pacheco, J.M., Santos, F.C.: Evolution of collective action in adaptive social structures. Sci. Rep. **3**, 1521 (2013)
7. Mui, L.: Computational models of trust and reputation: agents, evolutionary games, and social networks. Ph.D. thesis, Massachusetts Institute of Technology (2002)
8. Ostrom, E.: Understanding Institutional Diversity. Princeton University Press, Princeton (2005)
9. Ostrom, E.: Design principles of robust property-rights institutions: what have we learned. In: Ingram, G.K., Hong, Y.H. (eds.) Property Rights and Land Policies: Proceedings of the 2008 Land Policy Conference. Lincoln Institute of Land Policy, Cambridge (Massachusetts) (2009). https://www.lincolninst.edu/sites/default/files/pubfiles/property-rights-andland-policies-chp.pdf
10. Rogge, N., Theesfeld, I., Strassner, C.: Social sustainability through social interaction-A national survey on community gardens in Germany. Sustainability **10**(4), 1–18 (2018). https://doi.org/10.3390/su10041085
11. Van Dam, K.H., Nikolic, I., Lukszo, Z.: Agent-Based Modelling of Socio-technical Systems, vol. 9. Springer Science & Business Media, Dordrecht (2012)

Chapter 16
Unleashing the Agents: From a Descriptive to an Explanatory Perspective in Agent-Based Modelling

Christopher K. Frantz

Abstract Agent-based modelling endows the experimenter with high levels of flexibility and, consequently, responsibility. Possibly because of that, developing good models is hard. In this work, we engage in the discussion around improving the analytical value and disciplinary acceptance of agent-based social simulation. To this end, this paper includes the proposal to make the agents themselves observers, as opposed to just participants, of the simulation to introduce explanatory power that cannot be leveraged by on descriptive macro-level analysis alone. This is followed by an argument for the use of institutional concepts for any mechanism that seeks to embed quasi-reflective capabilities in an effort to gain accessible explanatory insights from simulations. To exemplify this idea, we apply it to a cooperation game of moderate complexity and finally discuss application opportunities, challenges and future directions.

Keywords Agent-based modelling · Social simulation · Methodology · Explanatory simulation · Institutions · Norms · Corruption game · Institutional analysis · Grammar of institutions · Nested ADICO · Stereotyping · Implicit social cognition · Social learning · Institutional Modelling · nADICO

16.1 Introduction

Agent-based modelling and simulation (ABM) [22] is experiencing uptake for an increasingly wide range of coordination and cooperation problems based on its accessible agent metaphor and the ability to reconstruct problems incrementally and from the perspective of the problem domain. A specific opportunity arising from the application of ABM is its ability to inform modelling from a theory-driven perspective and/or based on existing empirical data [49], making ABM suitable

C. K. Frantz (✉)
Norwegian University of Science and Technology (NTNU), Gjøvik, Norway
e-mail: cf@christopherfrantz.org; christopher.frantz@ntnu.no

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_16

both for abstract conceptual work and for concrete applications (e.g. simulating behaviour during emergencies [42]).

However, on the flip side, the flexibility of the agent concept can be problematic. On the one hand, the principles of agent-based modelling encourage experimenters to think in terms of the problem domain and do not constrain them to selectively favour complexity of the problem domain or the embedded agent concept. On the other hand, the flexibility of the agent concept allows for the encoding of agent behaviour on arbitrary levels of complexity, construed either as simple execution rules or as complex architectures that are able to account for cognitive and social determinants of behaviour.

To manage the complexity of the resulting scenarios, ABM makes it convenient for experimenters to detach themselves from the underlying agent implementation, and rather *ascribe* agency (and potentially intentionality) to the modelled entities, and focus on the analysed (generally macro-level) phenomenon without taking agency on the micro level into account. Combined with the analytical focus on the problem domain, this leaves researchers at risk to perceive the underlying agent model (if not the entire simulation) as a black box and treat the produced results at face value during the ensuing interpretation.

The resulting inability to account for the introduced assumptions and abstractions, alongside other methodological concerns (which we discuss in Sect. 16.2), challenges the broader adoption of ABM in various disciplines and is frequently put forth in advocacy for methods that can rely on a comprehensive formalisation of the problem.

In this work, we explore how we can address this concern and shift the agent itself back into the spotlight as a means to provide better explanatory insights into the model dynamics. To achieve this, we will, of course, need to introduce a necessary minimal set of assumptions about the agents, which is compatible with calls to endow agent-based models with stronger social-psychological capabilities (see e.g. [27]). To illustrate this aspect, we will explore this idea using a conceptual cooperation problem with moderate socio-structural complexity that emulates prototypical behaviours found in economic exchange, such as corruption.

The paper is structured as follows: Sect. 16.2 sets the scene that motivates the use of agent conceptions that exhibit explanatory functions (while driving the phenomenon of interest). In Sect. 16.3, we develop a candidate approach to leverage deeper insights into the dynamics of agent-based models. Section 16.4 sketches a cooperation scenario that employs the proposed architecture, which is subsequently evaluated in Sect. 16.5. Section 16.6 concludes the paper with a discussion of the insights and outlines further research directions.

16.2 Background

The principles of social simulation more generally, and agent-based modelling specifically, have come a long way. Since Schelling's experimentation with cellular automata to analyse sociological phenomena [44] – marking the birth of social simulation – Axelrod's seminal work on cooperation [2] shifted agent-based concepts into the mainstream, and Epstein's declaration of simulation as the 'third way of doing science' [14] marked the methodological rite of passage. The accessibility of the intuitions underlying the agent concept, the availability of de facto standard modelling platforms,[1] and increasing maturity of methodological prescriptions and documentation standards (e.g. ODD + D [36]) lowered the threshold for the use of agent-based simulation in a wide range of disciplines, including political science [9], economics [15], institutional analysis [18], social psychology [26], criminology [6] and religious violence [46], to name a few.

Beyond the use as a tool for the analysis of specific phenomena, the principles of agent-based modelling have contributed to the exploration of fundamental sociological concepts, such as the role of trust for cooperation, social functions of reciprocity and the influence of topology on opinion formation and dynamics.[2]

However, confronted with the complexity of interaction on a social level, modellers are required to make strong assumptions about the underlying agent concept, as reflected in the KISS vs. KIDS discussion [12]. This includes the decision whether to model agents as primitive rule executors without any autonomy and prescribed social interaction or to opt for richer agent architectures that account for cognitive and social capabilities of humans,[3] the consideration of bounded rationality [47] or the scenario-dependent situational adaptation of behavioural strategies (e.g. Janssen and Jager [28]). At the same time, it is at the modeller's discretion to decide how interactions between individuals are represented (e.g. in comprehensive detail or as compound action) and to what extent agents can observe their physical and social environment (limited observation, noise,[4]) as well as their ability to retain and access information.

This flexibility, the seemingly arbitrary choice of detail and subsumption of socio-cognitive functions by abstract architectures, made ABM subject to criticism, including the objection to high-level abstractions, choice of assumptions and their empirical support [31], epistemological challenges in identifying causal relationships in the first place [24],[5] practitioners' concerns with the abstract representation of agency [32] and, last but not least, challenges from a methodological standpoint [20, 21].

Despite these concerns, agent-based modelling provides conceptual riches and explorative potential for the analysis of social systems that few other techniques can offer. It builds on the human metaphor without carrying psychological burdens of

[1]See Kravari and Bassiliades [30] for a comprehensive overview. Abar et al. [1]'s survey provides a refined differentiation of platforms by application domains.

[2]Bianchi and Squazzoni [5] collated an insightful overview that illustrates the impact of ABM on sociology.

[3]For an overview refer to Balke and Gilbert [3].

[4]The importance of considering noise in the physical and social environment has been convincingly argued by Macy and Tsvetkova [34].

[5]Equally noteworthy is the rebuttal of Grüne-Yanoff's argument by Elsenbroich [13].

actual humans (biases, unintended learning effects, questionnaire fatigue, etc.), the control of which makes empirical studies with human participants expensive and error-prone. Being freed from such limitations, we propose to move beyond making agents mere actors in the scenarios of interest and exploit their psychologically impartial nature and deterministic properties to make the agents themselves quasi-reflective observers of the scenario. In doing so, we can endow agents with an *explanatory* role following the motto: “Don’t tell me *what* you do, tell me *why* you do it.”

However, before developing this proposal in greater detail in the following section, it is important to guard against potential misconceptions of the proposed approach.[6] The seasoned modeller may suggest that most agent-based modelling platforms, in fact, offer mechanisms that allow the runtime inspection of agent properties – an aspect that relates to the intuitions of this work.[7] However, while such functionality exists, it is (a) generally intended to support the development process in order to debug agent properties (e.g. resource levels) and (b) is focused on the situational state of the inspected entity. The approach put forth in the following sections qualitatively differs in that it provides richer statement representations that aim at reflecting the “narrative” of the scenario from the perspective of an agent – targeting the experimenter, as opposed to the developer. The proposed approach emphasises a dynamic perspective that captures and condenses the interaction history in an intuitively accessible syntactic form over the conventional comparative-static approach applied in the step-wise inspection of agent state, the latter of which leaves it to the experimenter to *manually infer* associated agent behaviour.

16.3 Concept

The central challenge in augmenting agents with human-like reflective capabilities – while retaining *scenario independence* of the approach and affording a *lightweight and accessible interpretation* – is to identify a basis to deliberate about the cognitive assumptions for a quasi-reflective agent.

16.3.1 Institution as a Cognitive Basis

Informing this decision, we could allude to the superior human reasoning capabilities and consequently favour concepts that emphasise deliberation abilities, such

[6]At this stage, it is important to acknowledge the anonymous reviewers who provided valuable feedback for further refinement.

[7]Noteworthy examples include Swarm [35], MASON [33], NetLogo [48] and Repast [40].

as represented by cognitive agent architectures. However, the focus on such risks misrepresenting the mechanisms that facilitate humans' functioning in social groups would neglect subconscious processes dominating routine-based decision-making (see, e.g. Kahneman [29]). Instead, for a realistic baseline representation of social functioning, let us suggest that we primarily rely on fundamental mechanisms that make our social environment computable by allowing us to develop predictive capabilities that accommodate the bounds of rationality [47], while being adaptive to changing social and situational circumstances – institutions. Institutions [25, 39], stylised as the "rules of the game" [38], are entrenched social behaviour, such as conventions (e.g. which side of the road to drive on), social norms (e.g. queueing for payment) and rules (e.g. traffic regulation, contracts) that are imposed by some authority or arise based on emergent behaviour (e.g. collective action) and are transmitted by socialisation. Essential characteristics for functioning institutions are their adoption, accepted normative status and subsequent embedding in participants' mental structure. This fundamental role of institutions becomes clearer when interpreting their establishment itself as self-referential, in that the "essence of belief is the establishment of habit" [43]. Searle [45] likewise deems institutional structures fundamentally embedded in our cognitive processes, and, assuming a more radical position, Castelfranchi suggests we can interpret "minds [themselves] as social institutions" [8].

In essence, if we assume that the belief in institutions (irrespective of the concrete form) is the lowest common denominator of any individual's (and, in extension, any society's) belief system, the use of institution representations is a sensible starting point for leveraging the explanatory power of agents.

While we briefly discussed the role of institutions as fundamental structure, we have yet to clarify the relevant processes that we assume for the associated agent model. One of those is the concept of "implicit social cognition" [23], visible in the ability to form and operate on observed patterns of individual and social characteristics – a specific function we commonly refer to as *stereotyping*. This implies the ability to draw generalisations across multiple attribute combinations, something we humans are specifically good at. More importantly, we are fast to do so [50] and willingly sacrifice accuracy and ignore representativeness. Another relevant function to understand and generalise social information is the ability to not only learn directly from personal experiences (*experiential learning*), but to learn from one's social environment by applying some form of *social learning* [4]. However, while we deem the ability to rely on stereotypes for heuristic purposes as essential for the processing of behavioural information, the ability to learn from the social environment introduces stronger assumptions about the agents' sensing abilities, and, in consequence, for simulation scenarios. It is for this reason that we consider this an optional component of such baseline architecture.

With this position in mind, we will turn to a candidate representation mechanism from the area of institutional modelling and analysis that allows us to integrate the fundamental processes described above.

16.3.2 Nested ADICO (nADICO) for Endogenous Inference of Social Institutions

When intending to provide a generic way to capture individuals' observations to infer its institutional function, we are, of course, subjected to a wide range of potential representation options, especially from the area of electronic institutions [37] and normative multi-agent systems [7]. Seeking for a generic cross-disciplinary approach, we employ a formalism that builds on Crawford and Ostrom [10, 11] *Grammar of Institutions*, borrowed from the area of institutional analysis [41]. The fundamental idea of the grammar is to rely on a uniform structure that allows the encoding of any form of institution (i.e. convention, norm or rule). For this purpose, the grammar consists of an *Attributes* component (A) that describes acting individuals' characteristics and a *Deontic* component (D) used to capture the normative signal as obligation, prohibition or permission. The actual action is encoded in the *Aim* component (I), and the activation conditions (such as location, time or previous actions) are represented in the *Conditions* component (C). Where existing, sanctions or consequences are specified in the *Or else* component (O). Using those components in varying combinations allows the capturing of different institution types. The combination of the AIC components is sufficient to express conventions (e.g. 'Drivers (A) drive (I) on the right side of the road (C).'). Social norms, in contrast, have a regulative character and include the deontic (ADIC) to describe the prohibition, permission or obligation attached to an expression (e.g. 'Drivers (A) *must* (D) drive (I) on the right side of the road (C).'). Rules, finally, exploit the entire structure (ADICO) by specifying a consequence for the expression's violation (e.g. 'Drivers (A) must (D) drive (I) on the right side of the road (C), *or else* they will be fined (O).').

While expressive in its ability to capture institutions, ADICO operates on the macro level, intended to analyse institutional outcomes in the context of institutional analysis. However, operationalising a representation that allows agents to *endogenously infer the normative function of observations at runtime* requires a refined structure, an aspect addressed by Nested ADICO (nADICO) [17, 19]. nADICO changes the semantics for normative specifications in observations (a) by allowing statements to retain information about consequences and other contextual information to substantiate the inferred understanding and (b) by allowing the combination and nesting of ADICO components to comprehensively capture actions, involved roles and actors, as well as associated normative content for both actions and consequences.

Using the rule example from above, this would translate into 'Drivers (A) must (D) drive (I) on the right side of the road (C), or else police officers (A) must (D) fine them (I) under any circumstances (C).', with the syntax `ADICADIC`. Other, more complex examples include the use of logical operators to describe the relationship between actions and consequences (e.g. `(ADIC and ADIC)ADIC` to represent the co-occurrence of actions and a single consequence; `ADIC(ADIC x/or ADIC)` to model both inclusive (`or`) and exclusive (`xor`) sanction alternatives, etc.). With those (very briefly described) refinements, this representation enables a comprehensive representation of complex behavioural traces.

The syntactic representation (*structure*) is augmented with a *process* that guides the aggregation and synthesis of observations into nADICO statements that represent an agent's normative understanding. As a first step, it involves the collection of observations under consideration of past actions, involved actors and received interaction feedback to generate institutional statements that reflect the observed behaviour. The ensuing multi-level generalisation of statements occurs based on observable non-unique social attributes (social markers, such as roles or occupation), and combinations thereof enable the representation of subjectively generalised behaviour patterns. A detailed specification of structural aspects and the norm inference process can be found in [19]; a comprehensive discussion of related literature and associated software can be found under https://christopherfrantz.org/nested-adico.

16.3.3 Intrusive vs. Non-intrusive Application

For its application, we differentiate between an *intrusive* and a *non-intrusive* approach, which determines the role of the discussed mechanism in the context of developed models. In both cases, agents act as observers and develop a normative understanding of the observed social and/or physical environment that is accessible to the experimenter. For the non-intrusive case, the extracted information is thus of explanatory value for the experiment observer, whereas in the intrusive case, the agent itself uses the collected information to inform its decision-making. In this case, the explanatory mechanisms thus become part of the analysed model itself.

This differentiation is essential for the flexible application of the proposed approach. While this approach is generic and "attachable" to existing agent models in the non-intrusive variant, using the proposed mechanism for agents' decision-making (i.e. feeding generalised information back into the simulation model – the intrusive application) would, of course, introduce an "ideological bias" with respect to the proposed cognitive model and associated capabilities.

16.4 Corruption Game

To explore this concept, we introduce an illustrative scenario that features complex interactions between different role-based actors and affords motivational autonomy of the agents based on experiential learning.

The scenario, which we refer to as the *Corruption Game*, borrows the structural characteristics of Axelrod's metanorm game [2] to inform action choices, but differs in that it ignores evolutionary aspects and refines the scenario (a) by explicitly

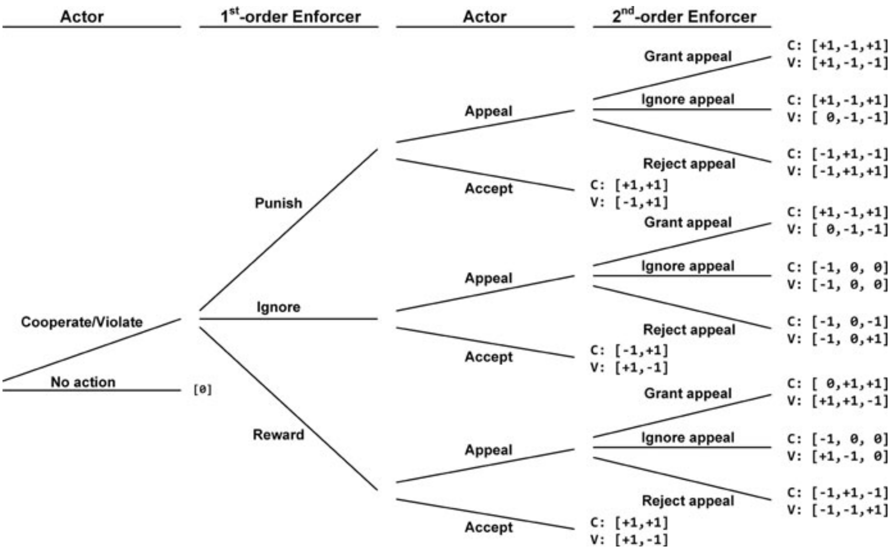


Fig. 16.1 Corruption game

modelling alternative action choices (including inaction) on the part of actors and enforcers and (b) by introducing explicit interaction of actor and second-order enforcer, aspects we will explore in detail in the following. The interaction schema of the game is depicted in Fig. 16.1.

The narrative underlying this game is the interaction of citizens with administrative officials that may react to transgressions (e.g. corruptive behaviour) by rewarding or punishing actors. Said officials are themselves subject to oversight by second-order officials who monitor their compliance as a response to citizens' complaints. This allows the exploration of prototypical scenarios, including administrative interactions such as handling tax returns or being punished for traffic violations, etc. – aspects that leave the first-order officials with considerable levels of discretion, making them potential participants in petty corruption. In the course of exploration, interesting questions revolve around the conditions under which the general behaviour shifts between violation and cooperation.

In the operationalisation, this translates into agents of two roles, either as citizens or officials (enforcers), with citizens pursuing cooperative or non-cooperative actions that are observed by enforcers, whose principal role is to reward cooperative behaviour and punish violations. As a third option, enforcers may simply ignore requests, which reflects institutional dysfunction, in contrast to wrongful decision-making by rewarding cheaters or punishing non-cheaters. Similarly, a citizen's inaction reflects the withdrawal from economic participation. Feedback associated with enacted action-reaction combinations is applied to all involved interaction partners and specified as part of the operationalisation.

Whatever the official's response, citizens can challenge any decision (or inaction) by appealing to a higher-level official, whose reaction determines the feedback for all involved stakeholders (citizen, first-order official, second-order official). Officials can act both as first- and second-order enforcers, but cannot act within the same transaction (i.e. an official cannot process the appeal against its own decision). The feedback for the action sequence chosen for this evaluation is denoted in Fig. 16.1 (Syntax: `[citizen[,1stOfficial[,2ndOfficial]]]`, with `1stOfficial` and `2ndOfficial` feedback only applying where interaction with officials takes place). For this baseline exploration, the feedback structure is modelled symmetrically (i.e. the extent of negative and positive feedback is identical) and rewards correctly identified cooperative behaviour, but equally rewards undetected cheating. The reason for this largely unbiased feedback specification is the exploration of social processes that mitigate or are decisive for the convergence towards violation or cooperative behaviour. In addition to introducing a more realistic breadth of action choices, the non-binary decisions are motivated by the ability of participants to withdraw from interactions if necessary – an aspect often ignored in analytical games, but a realistic indicator to assess the impact of corruption or institutional dysfunction.

As indicated in the conceptual description in Sect. 16.3, agents develop an understanding of normative behaviour by collecting experiential observations and aggregating feedback for generalised sequences of role attributes and associated actions. For this model, we will go beyond the generalisation of observations, and allow agents to inform their action choices using such observations (intrusive approach). Agents can thus retrieve the memorised feedback in aggregated form for different initial actions (e.g. violate) to drive their decision-making.

16.5 Evaluation

For the evaluation of the introduced model, we parametrised the scenario with the values shown in Table 16.1.

Table 16.1 Parameters

Parameter	Value range and step size
Number of citizens	25–75; step size: 25
Number of officials	25–75; step size: 25
Exploration probability	0.1
Cheater fraction	0.3–0.7; step size: 0.2
Cheating probability	0.5 (fixed)
Weight for observations	0.5 (fixed)
Memory length	100 (fixed)

Table 16.2 Correlation overview

Parameter	Cooperate	Violate	Inactive
Number of citizens	0.22	0.25	0.51
Number of officials	0.36	0.55	0
Quota of cheating citizens	−0.3	0.45	0
Social learning	−0.03	0.03	−0.25
Social learning separated by role	0.32	−0.22	−0.35
Ignoring actions	−0.38	0.36	0.51
Appealing	0.33	−0.14	−0.33

For the initial evaluation, we used the baseline scenario and selectively de/activated game characteristics (social learning,[8] ignoring actions, appealing) and systematically varied independent variables shown in Table 16.1 and measured the agents' preference for cooperative (COOPERATIVE) and deviant behaviour (VIOLATE), as well as for abstinence from any interaction (INACTIVE). The condensed results are shown in the correlation overview in Table 16.2.[9]

The results offer a mix of expected and interesting observations. With increasing number of citizens, we observe an increase in both cooperative and violation behaviour (with a mild tendency towards violations) but, more importantly, observe that actors increasingly abstain from participating in transactions. The variation of officials is likewise associated with compliance and violation but leads to stronger levels of violation behaviour. An increasing fraction of cheating citizens leads to an overall increase in violations, which is without surprise.

Social learning in itself does not have an impact on cooperative or violation behaviour. Instead, social learning appears to lead to an overall activation of participation. If limited to specific roles (i.e. citizen, official) – tagged as "social learning separated by role" – social learning leads to stronger levels of cooperative behaviour, along with an overall stronger activation of participants.

The final two parameters, selectively preventing agents from ignoring actions and appealing, have been introduced to reduce the game in breadth (ignoring) and depth (appealing) in order to understand the effects of those actions on cooperative behaviour.[10]

At this stage, we have reviewed initial results, and little choice but to take those at face value, making the interpretation prone to the problems described in Sect. 16.2, such as the oversimplified ascription of complex behaviour to agents,

[8]Social learning is operationalised as allowing agents to memorise fellow agents' institutional statement of the last action. For this operationalisation, the assumption is that all actions are overt. Agents' memory is bounded; they are able to store feedback for the last 100 experienced or observed interactions.

[9]We performed 5 runs for each parameter combination for 2000 rounds. All correlation values have been determined using Spearman's ρ.

[10]For the sake of focus, we will concentrate the discussion on the earlier parameters.

Table 16.3 Traces of citizen behaviour and correlation with citizen number

Statement	Correlation
CITIZEN: ACCEPT – OFFICIAL: SANCTION – CITIZEN: VIOLATE	0.58
CITIZEN: ACCEPT – OFFICIAL: REWARD – CITIZEN: VIOLATE	0.22
CITIZEN: ACCEPT – OFFICIAL: IGNORE – CITIZEN: COOPERATE	0.3
CITIZEN: ACCEPT – OFFICIAL: REWARD – CITIZEN: COOPERATE	0.25

and the inability to retrace the underlying processes. Following the motivation of this work, let us turn to the agents themselves and draw on their explanatory power to substantiate the insights. To achieve this, we discuss the impact of individual factors based on institutional statements recorded across all agents.

16.5.1 Citizen Numbers

Exploring the impact of citizen numbers, the initial observation in Table 16.2 is the stronger engagement both on cooperative and violation sides. Reviewing the relationship between number of citizens and prevalent statements in detail (see Table 16.3), we can make clarifying observations. Action choices resulting from an increasing number of actors have been absorbed into a few statements, here represented as simplified action sequences. These action sequences reflect generalised interaction patterns, with initiating actions noted on the right side and sequences building up to the left side. The first statement thus consists of three actions and posits that citizens accept an official's sanctioning after violating in the first place. Coming back to the specific results, in Statement 2, we can see that a side effect of the increase of citizens is an increase in mistaken rewards of violators by officials. This is insightful in that it indicates the extent to which citizens tolerate wrongful assessment and thus institutional dysfunction.

16.5.2 Social Learning

While our initial observations highlighted that social learning per se does neither favour cooperation nor violation, it supposedly leads to a stronger activation of participants. The statements that offer most insights (see Table 16.4) include the reduction in accepting an official's ignorance by cooperative citizens and reduction of citizens' inactivity, with an equal spread across other action variations (cooperation and violation).

Table 16.4 Traces of citizen behaviour and correlation with social learning

Statement	Correlation
CITIZEN: ACCEPT – OFFICIAL: IGNORE – CITIZEN: COOPERATE	−0.4
CITIZEN: IGNORE	−0.51

Table 16.5 Traces of citizen behaviour and correlation to role-separated social learning

Index	Statement	Correlation
1	OFFICIAL: REJECT_APPEAL – CITIZEN: APPEAL – OFFICIAL: SANCTION – CITIZEN: VIOLATE	0.38
2	CITIZEN: ACCEPT – OFFICIAL: IGNORE – CITIZEN: APPEAL – OFFICIAL: SANCTION – CITIZEN: VIOLATE	0.25
3	OFFICIAL: GRANT_APPEAL – CITIZEN: APPEAL – OFFICIAL: SANCTION – CITIZEN: VIOLATE	0.25
4	CITIZEN: ACCEPT – OFFICIAL: REWARD – CITIZEN: COOPERATE	0.38
5	CITIZEN: ACCEPT – OFFICIAL: IGNORE – CITIZEN: APPEAL – OFFICIAL: IGNORE – CITIZEN: COOPERATE	0.26
6	OFFICIAL: REJECT_APPEAL – CITIZEN: APPEAL – OFFICIAL: IGNORE – CITIZEN: COOPERATE	0.26
7	OFFICIAL: GRANT_APPEAL – CITIZEN: APPEAL – OFFICIAL: IGNORE – CITIZEN: APPEAL – OFFICIAL: SANCTION – CITIZEN: COOPERATE	0.13
8	OFFICIAL: REJECT_APPEAL – CITIZEN: APPEAL – OFFICIAL: SANCTION – CITIZEN: COOPERATE	0.26
9	OFFICIAL: GRANT_APPEAL – CITIZEN: APPEAL – OFFICIAL: SANCTION – CITIZEN: COOPERATE	0.23
10	CITIZEN: IGNORE	−0.35

16.5.3 Social Learning Separated by Role

While social learning promotes unbiased participation, when looking at role-separated social learning, the results point into a different direction. In this case agents only learn from their peers, which leads to a behavioural bias towards cooperative behaviour. How does this come about?

Looking at an excerpt of the collected statements (see Table 16.5), we can find a clue in the faster adoption of relevant information. By learning from their peers, agents quickly adopt preferable coordination behaviour. For example, agents are quick to learn that cooperative behaviour should be rewarded (Statement 4). However, exploring statements involving appeals processes offer stronger insights into the actual dynamics. As such, officials learn to reject appeals that are lodged by violators (Statement 1), but may also quickly adopt suboptimal behaviour, such as the granting appeals to violating citizens (Statement 3) and also learn that non-reaction to appeals (Statement 2) is a potential action alternative. Without discussing

```
A=A(*, {ROLE=[CITIZEN]}), D=3.0, I=I(APPEAL, *), C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[OFFICIAL]}), I=I(SANCTION, *),
        C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[CITIZEN]}), I=I(VIOLATE, *), C=C(*), O=(null)}), O=(null)}),
  O=(L1: A=A(*, {ROLE=[OFFICIAL]}), D=-3.0 (inv), I=I(GRANT_APPEAL, *), C=C(*), O=(null))

A=A(*, {ROLE=[CITIZEN]}), D=-1.0, I=I(APPEAL, *), C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[OFFICIAL]}), I=I(SANCTION, *),
        C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[CITIZEN]}), I=I(VIOLATE, *), C=C(*), O=(null)}), O=(null)}),
  O=(L1: A=A(*, {ROLE=[OFFICIAL]}), D=1.0 (inv), I=I(REJECT_APPEAL, *), C=C(*), O=(null))

A=A(*, {ROLE=[CITIZEN]}), D=-0.5, I=I(APPEAL, *), C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[OFFICIAL]}), I=I(REWARD, *),
        C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[CITIZEN]}), I=I(VIOLATE, *), C=C(*), O=(null)}), O=(null)}),
  O=(L1: A=A(*, {ROLE=[OFFICIAL]}), D=0.5 (inv), I=I(REJECT_APPEAL, *), C=C(*), O=(null))

A=A(*, {ROLE=[CITIZEN]}), D=0.5, I=I(APPEAL, *), C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[OFFICIAL]}), I=I(REWARD, *),
        C=C({PREVIOUS_ACTION=L0: A=A(*, {ROLE=[CITIZEN]}), I=I(COOPERATE, *), C=C(*), O=(null)}), O=(null)}),
  O=(L1: A=A(*, {ROLE=[OFFICIAL]}), D=-0.5 (inv), I=I(GRANT_APPEAL, *), C=C(*), O=(null))
```

Fig. 16.2 Excerpt of micro-level statements for action 'APPEAL'

all individual statements further, we can see a more refined dynamic that highlights stronger exploitation of complex institutional processes (i.e. utilising the depth of the action space).

16.5.4 Micro-level Inspections

While this approach allows us to retrace behavioural shifts in detail, we still operate on the macro level, based on aggregated preferred action choices of all involved agents, albeit at greater detail.

However, when exhausting the explanatory value on the macro level, the proposed approach allows us to drill deeper and explore individual agents' motivations for their behaviour. Agents develop conceptions of all explored and observed action choices, which is due to the generic nature of the approach, but it also offers a differentiated insight into the inner workings of agents and enables us, as experimenters, to assess where cognitive processes are sufficiently represented. Figure 16.2 shows an extract consisting of four statements of an agent's runtime understanding (in original syntax), centred around its decision-making with respect to the action "appeal". The statements clearly show how agents can operate with conflicting signals. The first statement, for example, suggests that an agent should appeal (positive deontic value) after an official's punishment of its violation. The motivation for this is reduced to the observation that chances are that the official may actually grant the appeal. The second statement effectively explores the opposing signal of discouraging appealing because of potential rejection. Similarly, the last two statements highlight conflicting motivations as to whether appealing after showing cooperative behaviour is useful.

Following this brief exposition, we can see that the mechanism is not only able to leverage explanatory insights into behavioural changes on the macro level, but also to transparently represent individual-level cognitive processes, such as the notion of cognitive dissonance [16] as just discussed above.

16.6 Discussion

In this paper, we argued for the use of a generic agent conception that satisfies a fundamental subset of processes found in social animals (stereotyping, social learning), and specifically humans, and attach or integrate this mechanism with existing agent-based models, so as to leverage these processes to provide additional explanatory power – in addition to the conventional aggregate macro-level observation of dependent variables. To explore model internals, agents collect and generalise their observations and represent those in a uniform way that allows their aggregation on arbitrary level of social organisation (e.g. individual observations, groups, or society at large). We showcased this approach using a moderately complex institutional scenario in order to explore the emerging behaviour at greater depth.

This work intends to drive the discussion around *exploring agency to understand agency* using institutional mechanisms. Given this motivation, the approach presented here is a candidate operationalisation in the form of a domain-independent baseline architecture that allows for the consideration of fundamental functions of human operation in social environments. Questions that invite for further discussion revolve around the minimal cognitive functions sufficient for a baseline operationalisation (here: stereotyping), and, if used as input for decision-making, in how far such architecture affords an "ideological buy-in" by experimenters and affects modelling freedom.

But returning to the motivation of this work, what are the concrete benefits of shifting from descriptive macro-level to explanatory micro-level approaches?

- Agents can be used as passive, non-intrusive observers, e.g. only used for verification of the model, or for the inspection of specific runs. The condensed generative conception of the institutional environment can thus be used for methodological support during model development.
- Using a generic institution operationalisation allows the detection of both intentional and unintentional behaviours (independent of the non-/intrusive application). This provides the experimenter with insight into both "desirable" and "undesirable" behaviours that may withdraw themselves from experimental observation for cases in which the underlying dynamics are obscured by aggregate metrics. This aspect is of analytical value, since it allows the experimenter to retrace explicit explanatory links between micro-level interaction dynamics and macro-level phenomena.

Both such aspects have the potential of contributing to building greater confidence in the development and analysis of agent-based models and, consequently, for the application of agent-based modelling more generally.

Future efforts involve the application of this approach to existing datasets to explore its usability for real-world applications, both in terms of usefulness and efficiency. This includes the intent to make this mechanism more readily available, e.g. as a plugin, to explore its value with new or existing simulation scenarios. Reflecting on the choice of an institution representation from the area of institutional analysis opens up manifold further interdisciplinary application opportunities. However, whatever the chosen mechanism, the essential argument this paper makes is that any approach to develop explanatory ABMs will, in one way or another, have to consider institutional concepts at its basis.

Concluding, we believe that it is important to drive a cross-disciplinary consensus that agent-based modelling has the capability to explore complex social phenomena but, unlike other quantitative approaches, can also offer ways to facilitate the interpretation of its own operation – and which vehicle would be more self-referential than the agents themselves?

References

1. Abar, S., Theodoropoulos, G.K., Lemarinier, P., O'Hare, G.M.: Agent based modelling and simulation tools: a review of the state-of-art software. Comput. Sci. Rev. **24**, 13–33 (2017)
2. Axelrod, R.: An evolutionary approach to norms. Am. Polit. Sci. Rev. **80**(4), 1095–1111 (1986)
3. Balke, T., Gilbert, N.: How do agents make decisions? A survey. J. Artif. Soc. Soc. Simul. **17**(4), 13 (2014)
4. Bandura, A.: Social Learning Theory. General Learning Press, New York (1977)
5. Bianchi, F., Squazzoni, F.: Agent-based models in sociology. Wiley Interdiscip. Rev. Comput. Stat. **7**(4), 284–306 (2015)
6. Birks, D., Townsley, M., Stewart, A.: Generative explanations of crime: using simulation to test criminological theory. Criminology **50**(1), 221–254 (2012)
7. Boella, G., van der Torre, L., Verhagen, H.: Introduction to normative multi-agent systems. In: Boella, G., van der Torre, L., Verhagen, H. (eds.) Normative Multi-agent Systems, Dagstuhl Seminar Proceedings 07122, Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI). Schloss Dagstuhl (2007)
8. Castelfranchi, C.: Minds as social institutions. Phenomenol. Cogn. Sci. **13**(1), 121–143 (2014)
9. Cederman, L.: Computational models of social forms: advancing generative process theory. Am. J. Sociol. **110**(4), 864–893 (2005)
10. Crawford, S.E., Ostrom, E.: A grammar of institutions. Am. Polit. Sci. Rev. **89**(3), 582–600 (1995)
11. Crawford, S.E., Ostrom, E.: A grammar of institutions. In: Understanding Institutional Diversity, chapter 5, pp. 137–174. Princeton University Press, Princeton (2005)
12. Edmonds, B., Moss, S.: From KISS to KIDS – an 'anti-simplistic' modelling approach. In: Davidsson, P., Logan, B., Takadama, K. (eds.) Multi-agent and Multi-agent-Based Simulation. Lecture Notes in Computer Science, vol. 3415, pp. 130–144. Springer, Berlin (2005)
13. Elsenbroich, C.: Explanation in agent-based modelling: functions, causality or mechanisms? J. Artif. Soc. Soc. Simul. **15**(3), 1 (2012)
14. Epstein, J.M.: Agent-based computational models and generative social science. Complexity **4**, 41–60 (1999)
15. Farmer, J.D., Foley, D.: The economy needs agent-based modelling. Nature **460**, 685–686 (2009)
16. Festinger, L.: A Theory of Cognitive Dissonance. Stanford University Press, Palo Alto (1957)
17. Frantz, C., Purvis, M.K., Nowostawski, M., Savarimuthu, B.T.R.: nADICO: a nested grammar of institutions. In: Boella, G., Elkind, E., Savarimuthu, B.T.R., Dignum, F., Purvis, M.K. (eds.) PRIMA 2013: Principles and Practice of Multi-agent Systems. Lecture Notes in Artificial Intelligence, vol. 8291, pp. 429–436. Springer, Berlin (2013)
18. Frantz, C., Purvis, M.K., Nowostawski, M.: Agent-based modeling of information transmission in early historic trading. Soc. Sci. Comput. Rev. **32**(3), 393–416 (2014)

19. Frantz, C.K., Purvis, M.K., Savarimuthu, B.T.R., Nowostawski, M.: Modelling dynamic normative understanding in agent societies. Scalable Comput. Pract. Exp. **16**(4), 355–378 (2015)
20. Galán, J.M., Izquierdo, L.R.: Appearances can be deceiving: lessons learned re-implementing Axelrod's 'evolutionary approach to norms'. J. Artif. Soc. Soc. Simul. **8**(3), 2 (2005)
21. Galán, J.M., Izquierdo, L.R., Izquierdo, S.S., Santos, J.I., del Olmo, R., López-Paredes, A.: Checking Simulations: Detecting and Avoiding Errors and Artefacts, pp. 119–140. Springer International Publishing, Cham (2017)
22. Gilbert, N.: Agent-Based Models. Sage Publications, London (2008)
23. Greenwald, A.G., Banaji, M.R., Rudman, L.A., Farnham, S.D., Nosek, B.A., Mellott, D.S.: A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. Psychol. Rev. **109**(1), 3–25 (2002)
24. Grüne-Yanoff, T.: The explanatory potential of artificial societies. Synthese **169**(3), 539–555 (2009)
25. Hodgson, G.M.: What are institutions? J. Econ. Issues **40**(1), 1–25 (2006)
26. Jackson, J.C., Rand, D., Lewis, K., Norton, M.I., Gray, K.: Agent-based modeling: a guide for social psychologists. Soc. Psychol. Personal. Sci. **8**(4), 387–395 (2017)
27. Jager, W.: Enhancing the realism of simulation (EROS): on implementing and developing psychological theory in social simulation. J. Artif. Soc. Soc. Simul. **20**(3), 14 (2017)
28. Janssen, M., Jager, W.: An integrated approach to simulating behavioural processes: a case study of the lock-in of consumption patterns. J. Artif. Soc. Soc. Simul. **2**(2) (1999). http://jasss.soc.surrey.ac.uk/2/2/2.html
29. Kahneman, D.: Thinking, Fast and Slow. Farrar, Straus and Giroux, New York (2013)
30. Kravari, K., Bassiliades, N.: A survey of agent platforms. J. Artif. Soc. Soc. Simul. **18**(1), 11 (2015)
31. Lengnick, M.: Agent-based macroeconomics: a baseline model. J. Econ. Behav. Organ. **86**, 102–120 (2013)
32. Levy, S., Martens, K., van der Heijden, R.: Agent-based models and self-organisation: addressing common criticisms and the role of agent-based modelling in urban planning. Town Plan. Rev. **87**(3), 321–338 (2016)
33. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., Balan, G.: MASON: a multiagent simulation environment. Simulation **81**(7), 517–527 (2005)
34. Macy, M., Tsvetkova, M.: The signal importance of noise. Sociol. Methods Res. **44**(2), 306–328 (2015)
35. Minar, N., Burkhart, R., Langton, C., et al.: The swarm simulation system: a toolkit for building multi-agent simulations. Technical report (1996)
36. Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., Schwarz, N.: Describing human decisions in agent-based models odd+d, an extension of the odd protocol. Environ. Model. Softw. **48**, 37–48 (2013)
37. Noriega, P.: Agent-mediated auctions: the Fishmarket Metaphor. Ph.D. thesis, Universitat Autonoma de Barcelona, Barcelona (1997)
38. North, D.C.: Institutions, Institutional Change, and Economic Performance. Cambridge University Press, New York (1990)
39. North, D.C.: Institutions. J. Econ. Perspect. **5**(1), 97–112 (1991)
40. North, M.J., Collier, N.T., Ozik, J., Tatara, E., Altaweel, M., Macal, C.M., Bragen, M., Sydelko, P.: Complex adaptive systems modeling with Repast Simphony. Complex Adapt. Syst. Model. **1**(1), 3 (2013)
41. Ostrom, E.: Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge University Press, New York (1990)
42. Pan, X., Han, C.S., Dauber, K., Law, K.H.: A multi-agent based framework for the simulation of human and social behaviors during emergency evacuations. AI Soc. **22**(2), 113–132 (2007)
43. Peirce, C.S.: How to make our ideas clear. Pop. Sci. Mon. **12**, 286–302 (1878)
44. Schelling, T.C.: Dynamic models of segregation. J. Math. Sociol. **1**(2), 143–186 (1971)
45. Searle, J.R.: What is an institution? J. Inst. Econ. **1**(1), 1–22 (2005)

46. Shults, F.L., Gore, R., Wildman, W.J., Lynch, C., Lane, J.E., Toft, M.: Mutually escalating religious violence: a generative model. In: Proceedings of the 2017 Social Simulation Conference, Dublin (2017)
47. Simon, H.A.: A behavioral model of rational choice. Q. J. Econ. **69**(1), 99–118 (1955)
48. Tisue, S., Wilensky, U.: NetLogo: design and implementation of a multi-agent modeling environment. In: SwarmFest, Ann Arbor (2004)
49. Tolk, A.: Learning Something Right from Models That Are Wrong: Epistemology of Simulation, pp. 87–106. Springer International Publishing, Cham (2015)
50. Zeithamova, D., Schlichting, M., Preston, A.: The hippocampus and inferential reasoning: building memories to navigate future decisions. Front. Hum. Neurosci. **6**, 70 (2012)

Chapter 17
Participatory Policy Development with Agent-Based Modelling: Overcoming the Building Energy-Efficiency Gap

Eva Halwachs, Anne von Streit, and Christof Knoeri

Abstract To support the energy transition, making buildings, and in particular residential buildings, more energy-efficient is a central issue. To develop effective policies for achieving an increase of the renovation rate, regional and communal initiatives and policy structures play an important role. Decisions regarding energy-efficient renovations are characterized by the local interaction of homeowners, construction experts and regulators with the physical properties of the building stock. Agent-based modelling is an ideal tool, to display such complex interplay of socio-technical systems and the interacting actors. Regional-specific models can therefore support regional decision-makers in the policy development process. This paper describes (i) how agent-based modelling can contribute to a common system understanding by simulating different pathways of the regional building stock and its future energy demand and (ii) how social simulation can support policy actors in the development of regional policies concerning buildings' energy efficiency.

Keywords Participation · Policy · Building stock model · Energy efficiency

17.1 Introduction

Energy demand from buildings accounts for about 35% of the total energy demand globally [1]. Increasing the energy efficiency of the building stock, especially the residential building stock, is a critical component of national strategies to

E. Halwachs (✉) · A. von Streit
Human-Environment Relations, Department of Geography, Ludwig Maximilian University of Munich, Munich, Germany
e-mail: eva.halwachs@lmu.de

C. Knoeri
Group for Sustainability and Technology, Department of Management, Technology and Economics, ETH Zurich, Zürich, Switzerland

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_17

reach climate change targets (e.g. Germany plans to reduce the energy demand in the housing sector by 80% until 2050 [2]). To achieve these national targets, it is essential to increase the annual renovation rate [3]. Though a wide range of initiatives to increase the renovation rate exist in Germany and the total economic benefits of retrofitting usually outweigh the costs [4], the current renovation rate is still incredibly low [5]. Recent studies show that the current policy mix is not fully effective to address the barriers in homeowners' renovation decision-making and that noneconomic factors are highly underrated [6, 7]. Due to the complex interplay of social, technical and regulatory factors and the heterogeneity of actors and their local interaction, an integrative approach is needed.

To support the sustainable energy transition, regional and communal policies have gained an important role in the past years. Regional initiatives, aiming at energy autarky or being independent from fossil fuels, have contributed to developing new institutional structures [8, 9]. Energy consulting and information campaigns on the regional level have a better understanding of regional and local characteristics and have an easier access to important target groups. Policies aiming at specific target groups (e.g. building or household characteristics) can therefore be implemented more easily. Also, due to the municipal planning sovereignty in most European countries, local governments can influence the implementation of and compliance with building regulations.

The interplay of policies and individual behaviour (e.g. household) and the effects on regional building energy demand is a challenge regional governments are currently facing. Agent-based modelling (ABM) can display this interaction of policies and individual behaviour [10] and has in recent years already been used successfully to simulate the effect of policies on energy-efficient renovation and the building's subsequent energy demand [11, 12]. Also, participatory modelling has in the meanwhile become a common tool for planning processes and collective decision-making in various disciplines [13–15]. However, using ABM as a communication tool and participatory instrument in processes with political actors is not yet an established practice.

ABM provides a natural description of a system and is easily adapted to current challenges [16] and therefore offers an ideal basis for participatory processes [17, 18]. We therefore propose to develop an agent-based model that supports regional decision-makers in developing effective policies for the regional energy transition. The model is used as an instrument for communication in a participatory process with regional stakeholders to (i) present and discuss potential energy saving through renovation and heating system change and (ii) to test regional and communal policies developed in a participatory process and simulate their effects on household's renovation behaviour.

In this paper we briefly describe the case study region, outline the agent-based model and participatory process, present first results and elaborate on their implications.

17.2 Case Study

This study is embedded in the transdisciplinary project INOLA[1], which aims at supporting the transition process to a sustainable energy system in three counties in Upper Bavaria, Germany. All three counties – Bad Tölz-Wolfratshausen, Miesbach and Weilheim-Schongau – are economically prospering and expected to experience population growth, which is likely to lead to a higher energy demand in the future [19].

The study region set itself the target to become independent from fossil energy[2] by 2035. However, currently only 13.8% of the region's heat demand is covered by renewable energy [20]. Therefore, the reduction of buildings' energy demand is an essential component of the strategy to achieve the regional energy transition.

17.3 Methods

17.3.1 Agent-Based Model

The purpose of the agent-based model is to simulate households' renovation behaviour and its impact on regional building energy demand over time. The model has been developed in two stages. (i) First, a basic bottom-up building stock model based on [3, 12] was created solely to demonstrate the broad effects of energy-efficient renovation on regional energy demand. The model portrays the regional energy demand based on building-relevant characteristics (e.g. building type, age and heating type) and the development over time. These changes are dependent on the set of parameters (e.g. renovation rate, construction rate, renovation standard and new building standard). The model setup was validated on the system level with consumption data provided by regional and communal energy utilities. (ii) After a first workshop (see Sect. 17.3.2), this basic probabilistic building stock model has been extended with a behaviour model based on a survey about household renovation decision-making (see Fig.17.1). By including empirical data, more 'behaviourally realistic agents' [21] can be modelled.

The survey included personal as well as building-related factors, actors involved in the renovation decision and contextual factors (e.g. interest rate). By using such detailed empirical data as model basis, it is possible to test policies targeting at specific households or buildings. These policy measures might, for example, include a ban on certain energy carriers when renovating or building a house, incentives

[1]The project INOLA (Innovations for a Sustainable Land Use and Energy Management on a Regional Level; http://innovationsgruppen-landmanagement.de/en/innovationsgruppen/ inola/) is funded from 2014 to 2019 within the Framework Programme FONA (Research for Sustainable Development) by the German Federal Ministry of Education and Research.

[2]Excluding transport.

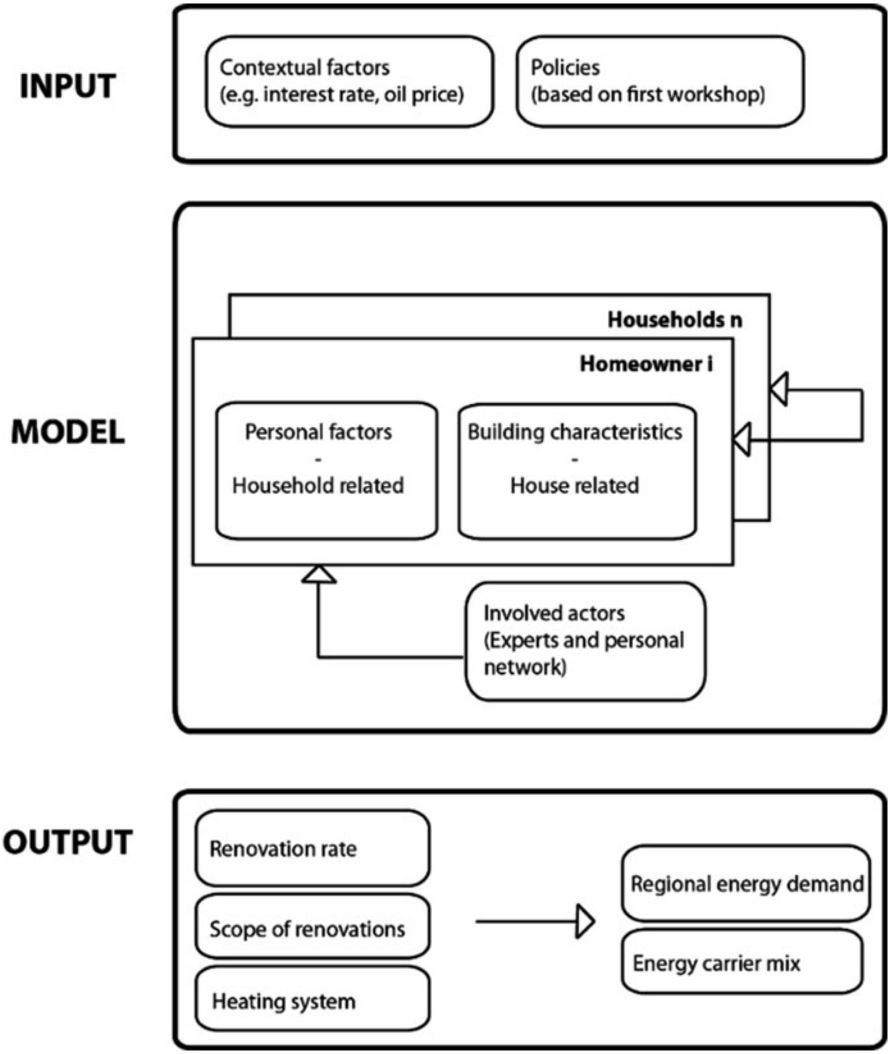


Fig. 17.1 Model overview

for the application of certain technologies on a municipality level and insulation campaigns addressing households with certain characteristics or living in certain building types. Additional outcomes to the basic model are the annual renovation rate, the scope of renovation measures and more detailed information about heating systems.

Table 17.1 Basic building stock scenarios

Scenario	Description
Business as usual	Continuation of past 10 years
Renovation	Increasing the renovation rate to the national aim of 2%
Legislation 1	Tightening of renovation standards
Legislation 2	Ban on new oil heating systems
Subsidies	Financial incentives on wood-based and heat-pump heating systems

17.3.2 Participatory Process

Policy scenarios are developed in a two-step participatory process, with two regional workshops to be conducted. This approach was used to first introduce the concept of ABM in general and to provide initial basic results as a start for discussion. Participants of all workshops are regional policy makers, energy advisors and other regional stakeholders from business or politics and the broader public.

17.3.2.1 Step 1

The aim of the first regional workshop is (i) to present the basic probabilistic building stock model and its functioning, (ii) to discuss the untapped potential of buildings' energy-efficiency improvements in the region and (iii) to discuss possible regional-specific policy measures to address this potential.

Based on the basic probabilistic model, five different scenarios (see Table 17.1) representing different policy actions were prepared for the workshop and presented in the plenum. These scenarios display the potential energy savings and were the basis for the following discussion: How can the region achieve the reduction of buildings' energy demand? What are potential regional policy instruments? How can existing policy instruments be improved?

17.3.2.2 Step 2

The second and upcoming regional workshop is designed to discuss the impact of the specific policy instruments which have been developed in the first workshop. These instruments will be implemented in the behavioural model to test the effects of each policy. These more detailed results of the extended model are theorised to support the decision and consensus finding on different policy mixes. The desired output is consensus on feasible regional measures to be tested in a pilot scheme in one municipality of the region.

17.4 Preliminary Results

First results based on the basic probabilistic model show that an increase of the renovation rate will achieve the highest annual energy savings by 2035. A focus on the renovation rate is therefore needed. Strengthening the energy standard of newly built houses had only little effect on the regional energy demand. Therefore, higher standards for new buildings only are not sufficient to achieve the energy transition aims. Highest reduction in the buildings' energy demand per annum (comparing 2015 and 2035) is achieved when combining a high renovation rate with high renovation standards. However, there might be a trade-off between the two measures: for example, a renovation standard that is too high might reduce the rate of energy-efficient renovations which can be tested with the extended behaviour model.

Policy makers in the region therefore need to discuss how the annual renovation rate can be increased without losing out on energy-efficiency and how more people can be reached through different sets of policies. By presenting the results of the basic scenarios in the first workshop, a common system understanding between researchers, policy decision-makers and broad public was enhanced. This helped to discuss further possibilities of energy-efficiency policies targeted at the regional energy demand, for example, the importance of renovating buildings with older construction years.

The next step is the supplementation of the behavioural model with the empirical survey data and output from the participatory workshop. Through that the ABM can be significantly improved and is likely to offer further possibilities for regional policy testing.

Funding statement: Financial support from the German Ministry of Federal Ministry of Education and Research (BMBF) through grant number 033L155AN is gratefully acknowledged.

References

1. Patwardhan, A.P., Gomez-Echeverri, L., Nakićenović, N., et al. (eds.): Global Energy Assessment (GEA). Cambridge University Press, Cambridge (2012)
2. Bundesministerium für Wirtschaft und Energie.: Energieeffizienzstrategie Gebäude: Wege zu einem nahezu klimaneutralen Gebäudebestand (2015)
3. Binder, C.R., Knoeri, C., Hecher, M.: Modeling transition paths towards de-centralized regional energy autonomy: the role of legislation, technology adoption, and resource availability. Raumforsch. Raumordn. **74**(3), 273–284 (2016). https://doi.org/10.1007/s13147-016-0396-5
4. Chapman, R., Howden-Chapman, P., Viggers, H., et al.: Retrofitting houses with insulation: a cost-benefit analysis of a randomised community trial. J. Epidemiol. Community Health. **63**(4), 271–277 (2009). https://doi.org/10.1136/jech.2007.070037
5. Deutsche Energie-Agentur GmbH.: Der dena-Gebäudereport 2012: Statistiken und Analysen zur Energieeffizienz im Gebäudebestand (2012)

6. Weiss, J., Dunkelberg, E., Vogelpohl, T.: Improving policy instruments to better tap into homeowner refurbishment potential: lessons learned from a case study in Germany. Energy Policy. **44**, 406–415 (2012). https://doi.org/10.1016/j.enpol.2012.02.006
7. Friege, J., Chappin, E.: Modelling decisions on energy-efficient renovations: a review. Renew. Sust. Energ. Rev. **39**, 196–208 (2014). https://doi.org/10.1016/j.rser.2014.07.091
8. Hecher, M., Vilsmaier, U., Akhavan, R., et al.: An integrative analysis of energy transitions in energy regions: a case study of ökoEnergieland in Austria. Ecol. Econ. **121**, 40–53 (2016). https://doi.org/10.1016/j.ecolecon.2015.11.015
9. Späth, P., Rohracher, H.: 'Energy regions': the transformative power of regional discourses on socio-technical futures. Res. Policy. **39**(4), 449–458 (2010). https://doi.org/10.1016/j.respol.2010.01.017
10. Grimm, V., Railsback, S.F.: Individual-Based Modeling and Ecology. Princeton Series in Theoretical and Computational Biology. Princeton University Press, Princeton (2013)
11. Friege, J.: Increasing homeowners' insulation activity in Germany: an empirically grounded agent-based model analysis. Energ. Buildings. **128**, 756–771 (2016). https://doi.org/10.1016/j.enbuild.2016.07.042
12. Knoeri, C., Goetz, A.: Generic bottom-up Generic bottom-up building-energy models for developing regional energy transition scenarios. Paper presented at the 9th Conference of the European Social Simulation Association. Barcelona (2014)
13. Mendoza, G.A., Prabhu, R.: Combining participatory modeling and multi-criteria analysis for community-based forest management. For. Ecol. Manag. **207**(1–2), 145–156 (2005). https://doi.org/10.1016/j.foreco.2004.10.024
14. Gaddis, E.J.B., Falk, H.H., Ginger, C., et al.: Effectiveness of a participatory modeling effort to identify and advance community water resource goals in St. Albans, Vermont. Environ. Model Softw. **25**(11), 1428–1438 (2010). https://doi.org/10.1016/j.envsoft.2009.06.004
15. Röckmann, C., Ulrich, C., Dreyer, M., et al.: The added value of participatory modelling in fisheries management – what has been learnt? Mar. Policy. **36**(5), 1072–1085 (2012). https://doi.org/10.1016/j.marpol.2012.02.027
16. Bonabeau, E.: Agent-based modeling: methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. U. S. A. **99**(Suppl 3), 7280–7287 (2002). https://doi.org/10.1073/pnas.082080899
17. Johnson, P.: Agent-based models as "interested amateurs". Land. **4**(2), 281–299 (2015). https://doi.org/10.3390/land4020281
18. Gilbert, N., Ahrweiler, P., Barbrook-Johnson, P., et al.: Computational model-ling of public policy: reflections on practice. JASSS. **21**(1), (2018). https://doi.org/10.18564/jasss.3669
19. Böhme, S., Eigenhüller, L., Werner, D., et al.: Demografie und Arbeitsmarkt in Bayern: Entwicklung, aktuelle Lage und Ausblick. IAB-Regional - Berichte und Analysen aus dem Regionalen Foschungsnetz (2012)
20. Hofer, V., Süß, A., Prasch, M., et al.: Das naturräumliche und technische Potential für Erneuerbare Energien in der Modellregion Oberland. INOLA Working Paper (2016)
21. Knoeri, C., Binder, C.R., Althaus, H.: An agent operationalization approach for context specific agent-based modeling. JASSS. **14**(2), (2011). https://doi.org/10.18564/jasss.1729

Chapter 18
Go Big or Go Home? Simulating the Effect of Publishing Adopter Numbers for Two-Sided Platforms

Michelle D. Haurand and Christian Stummer

Abstract Two-sided digital platforms, such as eBay, enable the interaction between two sides of a market, for instance, people willing to sell (supply side) and people willing to buy (demand side) some item or service. Growth and success of these platforms are driven by cross-side network effects, as each side is attracted by a large number of participants on the other side and deterred by a low number. However, it is unclear whether and when it is beneficial or harmful to inform potential platform participants about the number of current adopters on the respective other side. We implement an agent-based simulation of platform development and use it to gain insights into the effect of publishing the number of adopters. In our sample application, we found a negative effect that is attenuated when publication of these numbers is postponed. While our results are limited in their generalizability, they show the potential of agent-based simulation as a useful tool for analyzing various strategies in the face of cross-side network effects for two-sided platforms.

Keywords Agent-based simulation · Two-sided platforms · Cross-side network effects

18.1 Introduction

Today, we live in what one might call the "platform age" [1], in which two-sided digital platforms have transformed value creation. Two-sided platforms can be classified into three types: market makers, audience makers, and demand coordinators, depending on whether they facilitate transactions (like selling old furniture), deliver messages (advertisements reaching potential customers), or "make goods and services that generate indirect network effects" [2]. Prominent platform examples include eBay, Amazon, and Airbnb (market makers), Facebook

M. D. Haurand (✉) · C. Stummer
Bielefeld University, Department of Business Administration and Economics, Bielefeld, Germany
e-mail: m.haurand@uni-bielefeld.de

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_18

(audience maker), and Visa and American Express (demand coordinators), all of which have been disruptive in their businesses [3] or have established themselves as valuable brands [4].

For two-sided platforms, common factors influencing adoption, such as word of mouth, normative social influence, and advertising, are particularly crucial, as their impact can be boosted through a special feature of two-sided platforms, namely, cross-side network effects (i.e., the number of adopters on one side of the platform influences the number of adopters on the other side) [5]. Still, there is a research gap with regard to quantifiable information on how cross-side network effects should be dealt with. Even without informing potential adopters about the number of platform members on the other side of the platform, there is an implicit cross-side network effect resulting from an increase in matches having a positive effect on word of mouth [6]. When explicitly informing them, this cross-side network effect might be even stronger.

So far, various platform operators appear to act quite differently in this respect. Market makers eBay and Amazon may serve as examples: while eBay publishes very recent numbers on the demand side, it publishes few about the supply side [7, 8]; Amazon does not actively advertise any numbers at all, leading to the public availability of only sparse and outdated information on the global demand side [9].

In the light of this discrepancy between strategies for publishing adopter numbers, this paper is concerned with two research questions: (1) What effect does publishing adopter numbers for two-sided platforms have on platform adoption on the other side? (2) How does this effect vary depending on the timing of releasing this information? To answer these questions, we employ an agent-based modeling approach, which is particularly valuable for studying the spreading of information and product diffusion [10]. We parameterize the model according to an existing market maker platform as a sample application.

This paper proceeds by introducing an agent-based simulation model of a two-sided market maker platform in Sect. 18.2. Section 18.3 then presents the results of our simulation, which are discussed in Sect. 18.4.

18.2 The Agent-Based Model

Our simulation model, which is based on prior work [6], takes into account heterogeneous agents generating either supply or demand. The two groups of agents only interact through an intermediating two-sided platform (see Fig. 18.1).

The upper arrows indicate actions of the (potential) adopters on either side of the platform with regard to themselves or the platform, while the lower arrows designate actions of platform operators with regard to (potential) adopters of either side of the platform. It is worth noting the slight asymmetry in the model, which results from the inherent differences between the supply side and the demand side. The agents on the supply side of a market maker platform are usually decision-makers in businesses, while the agents on the demand side are usually individuals who

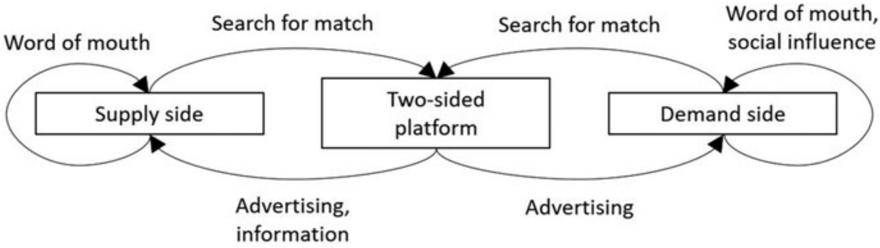


Fig. 18.1 Model of a two-sided platform joining supply and demand

are more socially interconnected and therefore likely to be influenced by observing the behavior of their peers and consequently imitate them in order to fit in. Thus, social influence is only present for the demand side. Still, agents of the supply and demand sides both search for a match from the other group through the platform, engage in word of mouth about it, and are subject to advertising measures by the platform operator. In our application case, word-of-mouth information is exchanged at regular conferences at which the (potential) adopters meet their peers. A more detailed description of agents and their behavior (e.g., with regard to the specifics of spreading information through word of mouth) in the basic model is provided in [6]. In the extended model at hand, we aim to study the effect of publishing information about the number of adopters of the platform. To this end, we mirror the real-life strategy of our sample application platform, that is, the platform operator only conveys information (if at all) about demand-side adopters targeted at the supply side (a strategy resembling that of eBay).

Our agents follow two straightforward rules concerning platform adoption: Firstly, they need to have been made aware of the platform's existence in order to consider it when looking for a match. Secondly, the number of agents they assume (or know, if they have received verified information) to be actually present on the platform must exceed an individual adoption threshold.

18.3 Results

Our simulation runs consider three scenarios: Firstly, the platform operator publishes no information about adopter numbers. Secondly, as was reality for our example platform, the platform operator starts conveying information about the number of demand-side agents on the platform to the supply-side agents in their advertising after the platform has existed for about 6 years and they have already acquired a certain user base. Thirdly, we simulate a scenario in which this number is communicated beginning with the foundation of the platform.

We carried out a face validation of the results with the platform operator, which is a highly recommended form of validation [11]. Furthermore, we performed

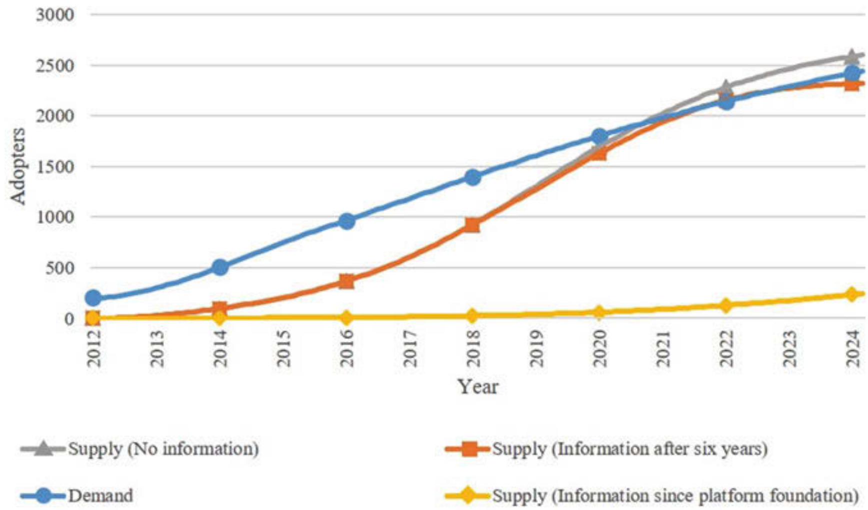


Fig. 18.2 Number of platform adopters over time for all three scenarios

extensive sensitivity analyses in order to ensure that our results are not compromised by the chosen parameter settings [12]. All past data could be verified by the platform operator to depict actual platform adoption.

The number of platform adopters on the supply side over time for all three scenarios (as well as the number of adopters on the demand side, which did not receive any information and thus did not vary significantly between scenarios) is depicted in Fig. 18.2.

18.4 Discussion

For the application case at hand, the optimal strategy with regard to platform growth turned out to be to not inform supply-side agents at all, as publishing adopter numbers had a negative effect on platform adoption. The extent of this effect depends on the timing of information release, with later provision of information being less detrimental to platform growth. It should be noted that this particular outcome should not be used to make generalizations and ultimately strongly depends on the adoption thresholds of the supply-side agents. While we could show the applicability of agent-based modeling for investigating the effects of publishing numbers of adopters for two-sided platforms, further research concerning the optimal timing of information publication and different adoption thresholds is still necessary. Moreover, we seek to validate simulation results by comparing it to the real-world development of other platforms.

References

1. Sussan, F., Acs, Z.J.: The digital entrepreneurial ecosystem. Small Bus. Econ. **49**(1), 55–73 (2017)
2. Evans, D.S.: The antitrust economics of multi-sided platform markets. Yale J. Regul. **20**(2), 325–381 (2003)
3. CNBC disruptor 50 list, https://www.cnbc.com/2017/05/16/the-2017-cnbc-disruptor-50-list-of-companies.html, last accessed 2019/03/14
4. Interbrand best global brands 2016 rankings, http://interbrand.com/best-brands/best-global-brands/2016/ranking/#, last accessed 2019/03/14
5. Dietl, H.M., Grütter, A., Lutzenberger, M.: Defensive competitive strategies in two-sided markets: the example of the mail industry. Compet. Regul. Netw. Ind. **10**(1), 3–15 (2009)
6. Haurand, M.D., Stummer, C.: The early-stage development of two-sided digital platforms: a simulation approach. In: Proceedings of ECIS 2018, Research Paper 49 (2018)
7. eBay, https://www.ebayinc.com/stories/news/ebay-inc-reports-fourth-quarter-and-full-year-2017-results/, last accessed 2019/03/14
8. eBay, https://www.ebayinc.com/stories/press-room/de/ebay-feiert-20-jahre-und-startet-ebay-plus-in-deutschland/, last accessed 2019/03/14
9. Statista, https://www.statista.com/statistics/476196/number-of-active-amazon-customer-accounts-quarter/, last accessed 2019/03/14
10. Kiesling, E., Günther, M., Stummer, C., Wakolbinger, L.M.: Agent-based simulation of innovation diffusion: a review. Cent. Eur. J. Oper. Res. **20**(2), 183–230 (2012)
11. Bonabeau, E.: Agent-based modeling: methods and techniques for simulating human systems. Proc. Natl. Acad. Sci. U. S. A. **99**(suppl 3), 7280–7287 (2002)
12. Ten Broeke, G., Van Voorn, G., Ligtenberg, A.: Which sensitivity analysis method should I use for my agent-based model? J. Artif. Soc. Soc. Simul. **19**(5), (2016)

Chapter 19
Simulations with Values

Samaneh Heidari, Maarten Jensen, and Frank Dignum

Abstract Using values as drivers of behavior has already been done in previous research. One of the most well-known universal theories of values is Schwartz's theory of abstract values. According to his theory, a universal set of abstract values can be imputed to people. As the values used in his system are very abstract, there is a need to translate the abstract values to more concrete values and assign the behavioral choices to them. A theory or methodology for this step has not been developed in a way that is widely applicable. Thus, a precise way of such a translation is necessary for practical purposes. In this paper, we design a practical but formal framework that can be used to study the value-driven behavior of agents in social simulations. We make an agent-based simulation for a fishery village that uses this framework.

Keywords Agent-based simulation · Policy-making · Policy values · Personal value · Value-based decision · Value framework

19.1 Introduction

The idea that values are abstract drivers of behavior is not new. What is interesting about the use of values is that, at least according to Schwartz [1], there is a universal set of abstract values that can be attributed to people. Differences between people stem not from having different values, but from giving different priorities to the values. This makes it possible to use values as a starting point to compare behaviors. The downside of the value theory of Schwartz is that the values defined

The first author has received funding from the European Union's Horizon 2020 Framework Programme Marie Sk lodowska-Curie (MSC) – ITN – ETN programme (project 642080).

S. Heidari (✉) · M. Jensen (✉) · F. Dignum (✉)
Utrecht University, Utrecht, Netherlands
e-mail: maarten.jensen@cs.umu.se; F.P.M.Dignum@uu.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_19

are very abstract and thus not directly related to behavior. Several steps are needed to translate abstract values into more concrete values and ultimately into behavioral choices. The way people concretize abstract values into concrete choices for action can also differ. Therefore, there is a need to describe this whole system in a precise and unambiguous way before it can be used for practical purposes. Some work of formally describing the relation between abstract values and actions, using values trees, has been done in [2]. In this paper we start from this logical framework and show how a quantitative framework can be designed that can be used to drive behavior of agents in social simulations.

A note should be made on the applicability of values as drivers of behavior. Not all behavior is primarily value driven. In normal life values usually play an explicit role only in larger (life changing) decisions, while smaller day-to-day behavior is governed by goals and norms. However, in many social simulations, we are exactly interested in situations where people do make life-changing decisions, such as moving houses, changing jobs, change for a more sustainable life style, etc. Thus it seems that the framework is relevant for many simulations.

In Sect. 19.2, we give some background on the value framework of Schwartz and the way we can connect these abstract values to concrete values and decisions on actions as described by Weide in [2]. In Sect. 19.3, we will discuss the framework that we propose to use to translate this theory into an implementable framework that can be used in agent-based social simulations. In Sect. 19.4, we illustrate how this framework can be used and leads to intuitive results in the domain of fishery management. Section 19.5 gives some conclusions and directions for future work.

19.2 Related Work

Schwartz et al. proposed ten basic values according to the universal needs of humans [1]. The Schwartz values are defined in the most abstract way that includes all the core values of every human all around the world.

As shown in Fig. 19.1, Schwartz value theory describes the dynamic compatibility and conflicting relation between all the value types by positioning them in a circle. Values close in the circle are more compatible than values on opposite sides. For example, pursuing $Tradition$ and $Hedonism$ are conflicting values, as $Tradition$ is about restraining owns actions to conform traditions and $Hedonism$ is about self-oriented need for pleasure. However, pursuing $Tradition$ value is compatible with pursuing $Conformity$ (to not violate social expectations in groups usually with close others) as both stress self-restraint and submission. In other words, the compatibility level of values in the Schwartz value circle decreases when the distance of them increases in the circle. The least distance belongs to two values next to each other, and the most distance belongs to two values that are on opposite position of each other.

Values and value systems such as from Schwartz have been used in many research efforts to explore the behavior of a complex system, studying human

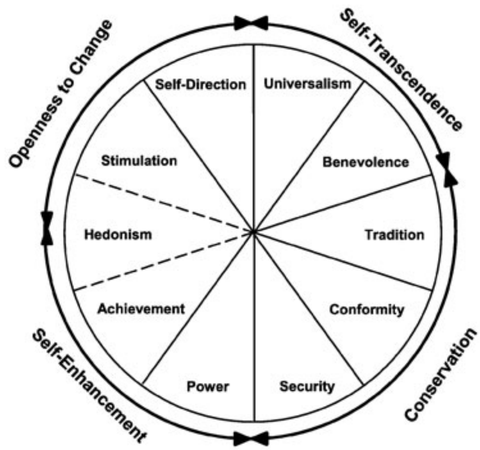


Fig. 19.1 Schwartz value circle, categorization and dynamicity of abstract personal values [1]

argumentation, managerial decisions, land-use behavior, and adaptation to climate change [3–6]. For example, the effects of individual values of a society on the general behavior of a complex system (including society, ecology, and economy) are studied in [7].

Bench-Capon et al.[8] show that promoting different values will lead to different arguments. Dechesne et al.[3] investigate how personal values (and other social phenomena) affect the behavior of people when introducing a smoking ban rule.

Mercuur [9] categorized usage of values in regard to doing an action into three main categories: pre-condition, post-condition, and deliberation navigator. In other words, values might be used as a measurement function to evaluate an action (post-condition), values can be used as a motivation to do an action (pre-condition) [10], and they can be used for justifying a decision of doing an action [2, 11].

Van der Weide [2] provides a formal model that can be used for modeling value-compliant decision-making. He shows how to form concrete values out of actions that can influence the abstract values. However, the relation between compatible and opposite values in the Schwartz value circle is not included in his model. Inspired by his formal model, we propose a framework for Schwartz value theory which not only considers translating Schwartz values into concrete values, but also the relation of values in the circle (Fig. 19.1), using values as pre-condition (filter) and justifying an action at the same time.

19.3 Framework

In this section, we propose a framework to make value-based decisions. Values can be used at different places in the deliberation cycle of the agents to select options (goals, plans, actions, etc.). If agents use goals and norms, then the values can be used to prioritize between those. Once goals are chosen and pursued, the values can be used to guide which plan is most in line with values. In this paper we focus on the motivational aspect of values which implies that they are at the basis of action selection. We start with the set of all salient actions (i.e., actions that can be taken at that moment because their pre-condition is true). If more actions are available the value tree and the current satisfaction of values are used to determine the highest priority of values. Then, the set of actions that are in line with the highest priority values will be chosen. From the resulting set of actions, one action is selected based on the current goal, the norms, and motives of the agent. In our current example, only social, economic, and ecological goals are used in this step. By performing an action, the agent updates his status. In what follows we explain the step in more detail.

In order to model the value circumplex of Schwartz, we define two sets. One of the input sets is a collection of Schwartz abstract values; $Values = \{V_1, V_2, V_3, V_4, V_5, V_6, V_7, V_8, V_9, V_{10}\}$, where $V_1 = Universalism$, $V_2 = Self\text{-}direction$, $V_3 = Stimulation$, $V_4 = Hedonism$, $V_5 = Achievement$, $V_6 = Power$, $V_7 = Security$, $V_8 = Tradition$, $V_9 = Conformity$, $V_{10} = Benevolence$. We need the indexes to consider the position of each value in the Schwartz circle in the framework.

The second set is the amount that each $V_i \in Values$ can get which is defined as $Importance = [1, 100]$. Any member $V_i \in Values$ can get any value from $Importance$ to identify its importance (which determines how often the value V_i has to be satisfied).

Assume that we define a function $\tau : Values \rightarrow Importance$ in which $\tau(V_i)$ gets the importance of value V_i. For each $V_i \in Values$, if $\tau(V_i) = 0$, then value V_i is silent and not playing a role in the system; if $\tau(V_i) = 100$, it is one of the important values if there are other values with the same importance.

To consider the relation of values, inducing the conflicting opposite values in the circle, we defined the following condition that shows how the importance of any members of $Values$ is related.

$$\text{Condition 1}: \forall i, j \in 1..10 : 0 \leqslant |\tau(V_\text{i}) - \tau(V_\text{j})| \leqslant \frac{|i' - j'|}{10} * 100,$$

where:

$$i' = \begin{cases} i & \text{if } 1 \leqslant i \leqslant 5 \\ 10 - i & \text{if } i > 5 \end{cases} \qquad j' = \begin{cases} j & \text{if } 1 \leqslant i \leqslant 5 \\ 10 - j & \text{if } j > 5 \end{cases}$$

in which 5 is the number of abstract values in one half of the Schwartz circle. Regarding symmetric distances of abstract values in the Schwartz circle, we slightly transform the formula by changing some variables.

Researchers used different version of Schwartz value system with various number abstract values. For example, Schwartz used seven abstract values to study the meaning of work in different cultures [12]. Also, it is possible to define different distance for items. In the current formula for Condition 1, we assumed the same area of each sectors. As an instance, the distance between $Universalism$ and $Tradition$ is the same as the distance between any other successive values in the circle. However, it is possible to change the formula to adapt different distances for items. For instance, according to this condition, it is possible to have this setting: $|\tau(Universalism) - \tau(Selfdirection)| = 15, |\tau(Power) - \tau(Achievement)| = 5$, and the distance of the other successive values remains 10. It would be the modelers preference to make such a decision according to their research requirements.

$$\text{Condition 2} : \begin{cases} \text{if } \tau(V_j) = 0 & \tau(V_i) > 50 \\ \text{if } \tau(V_j) \neq 0 & 0 \leqslant \tau(V_i) + \tau(V_j) \leqslant 100 \end{cases}, \text{where} j = (5 + i)\%10.$$

According to Condition 2, when value V_j is not included in the model ($\tau(V_j) = 0$), the opposite value of it in the Schwartz value circle should be high enough to have effects on the behavior of the system; otherwise, it can be ignored. All conditions can happen in extreme cases as well. For instance, it is possible to have all values with the same Importance. Also, having some values that do not play a role in the system ($\tau(V_i) = 0$) is possible.

In addition to the importance, values have level of satisfaction. In other words, people need to satisfy all of their values from time to time. But, the frequency of satisfaction differs due to their personal values. Function $\tau(V_i)$ shows how often value V_i should be satisfied. Therefore, there is a need to consider satisfaction level in the framework as well. To model changing in needs over time, we use the water tank model that determines the priority of satisfaction requirement for each value in Sect. 19.3.1.1.

19.3.1 Value-Based Selection

19.3.1.1 Value Satisfaction

In the Schwartz theory, all humans have ten introduced values, they consider their values in their life. Their life is consistent with their values. However, it is possible that in some conditions of the life, some values are not applicable. What makes a different personality is a different importance of values. As an example, consider a CEO of a multinational and an employee of an NGO. The NGO employee will do more activities that are in line with the $Universalism$ value, and the CEO will do more activities that satisfy $Power$. But, it does not mean that the NGO employee does not do any activity toward $Power$. The difference is the frequency and types of actions of satisfying the values. But, all the values need some level of satisfaction from time to time.

To model these dynamics, we use the water tank model represented in [13]. We consider one tank for each $V_i \in Values$. Each tank has the following base

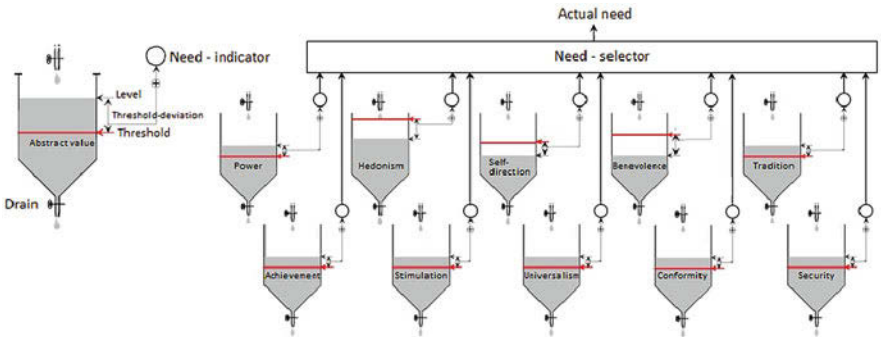


Fig. 19.2 Example figure of the water tank model for an agent

parameters: fluid level λ_i where $0 \leq \lambda_i \leq 100$ to indicate how much the value is satisfied and threshold $\tau(V_i)$ where $0 \leq \tau(V_i) \leq 100$ indicating when a value gets salient. The fluid drains every time step with a fixed amount of 10 to indicate that the value satisfaction is time dependent and increases when the agent does an action which is in line with the value. To be able to model the differing priorities of values of an agent, we use the threshold and calculate the priority. Agents try to fill up the tanks with the highest priority first. Priority of values is determined by using the following equation:

$$\rho = -((\lambda - \tau(V_i))/\tau(V_i)) * 100$$

We use negative sign as the priority of value satisfaction has reverse relation with its filled level. Filling up the tank can be done by performing actions that satisfy the abstract value connected to the tank. The increase amount is given by $(100 - \tau(V_i)) * \sigma$. This formula makes more important values fill up slower, thus increasing the frequency of performing actions relating to those values. It is possible to assign different values to multiplier σ for different actions. For example, buying a house usually has a larger effect on your values than buying an ice cream and thus has a larger multiplier σ for more impact.

A sample of the water tank model for an agent is shown in Fig. 19.2. Each agent has ten tanks, each with the same capacity and the same draining level. Though, the value thresholds are different for various agents.

19.3.1.2 Value Tree

The water tanks are used for determining which abstract value has to be satisfied. As mentioned prior, the abstract values do not directly impact the behavior of people, but rather through a series of perspectives that link the abstract values to concretized values that are directly related to behavioral choices. To make values work, we need to define more concrete values. Concrete values are easier to implement, and it is easier to track their contribution to a decision.

Several steps might be taken to translate an abstract value to its concrete values. One possible solution of formally describing the relation between abstract values and concrete values is through defining value trees [14]. The root node of the tree is an abstract value from Schwartz values. Nodes that are closest to the leaves are more concrete.

To view a simple example, one could look at Sect. 19.4 Fig. 19.4a in which the abstract values are $Power$, $Self-direction$, $Universalism$, and $Tradition$. The abstract values are the roots of the trees. As the values get more concrete, the further we go down from the root. Leaves of the trees are the most concrete values that related actions are assigned to them. By looking at the parent nodes of an action, we can determine which values it can satisfy and vice versa. Different path from each action to the root is deliberation that an agent uses to justify his action. For example, being a *fisher* as an action can satisfy $Tradition$ and/or $Universalism$.

People generally have different perspectives which can be modeled by giving only a subset of the total value tree to individuals. For example, to satisfy the $Universalism$ value through caring for the environment, agent A might buy an electric car because the emission of an electric car in use is less than a petrol car. Agent B might think electric cars are actually worse for the environment than petrol cars because of the chemicals used to create the batteries. He will instead go with public transport instead of his own car. This illustrates that two agents might perform different actions to satisfy the same abstract value. It can also be the case that agents perform the same action to satisfy two different abstract values. For example, playing a sport for one person can satisfy the $Achievement$ value (trying to win), while for the other, it satisfies the $Tradition$ value (play a game with friends as a way to be together). In other words, it is possible to assign different subsets of value trees to agents.

Some actions (and therefore their related concrete values) can be linked to more than one abstract value. Considering definitions of types of values introduced in [15], we can assign actions to abstract values for our case of interest, which is studying the behavior of a fishery village. For example, people in a fishery village might go fishing because they like nature ($Universalism$), they like adventure ($Stimulation$), they want to make money to promote their social status ($Power$), or they want to comply with their family traditional profession ($Conformity$).

Actions that are linked to compatible values might be positively interrelated. For example, actions that satisfy $Benevolence$ might have a positive effect on satisfying $Universalism$ as well. In contrast, if an action promotes a value, it can hardly attain the value opposite of it in the Schwartz circle.

19.3.1.3 Value-Based Filtering

Using values agents makes initial selections among the available actions to perform. We find the highest priority value that needs to be satisfied using the following formula:

$$\underset{V_i \in \text{Values}}{\arg\min}\ \rho(V_i) = \{V_i | V_i \in \text{Values}, \forall V_j \in \text{Values} : \rho(V_j) > \rho(V_i)\}$$

This formula returns the most preferred value (highest priority) in the current situation that needs to be satisfied. Then the actions promoting the highest priority value that are available are selected first. To compare the priority of each two values in order to find the highest priority, we use the following formula:

$$\forall V_i \in \text{Values}, V_j \in \text{Values} : \rho(V_i) = \rho(V_j) \text{ if } \rho(V_j) - \rho(V_i) < \delta$$

Meaning that $\rho(V_i)$ and $\rho(V_j)$ differ very little. Then all the actions that promote either V_j or V_i and are available get chosen.

The rules and conditions provided earlier are defined for abstract values in the Schwartz value system. All the concrete values in the value trees have the same importance as their root value. Therefore, all the rules and conditions of the abstract values (roots in value trees) are applicable to their related concrete values (leaves of the value trees).

It should be noted that it is possible to have some actions that are common between different value trees. For example, an agent can satisfy $Power$ or $Universalism$ by choosing to be a $Captain$ as $Captain$ is a shared action in these value trees (Fig. 19.3). However, the agent only satisfies one of the values by choosing action $Captain$ which depends on which deliberation he did before

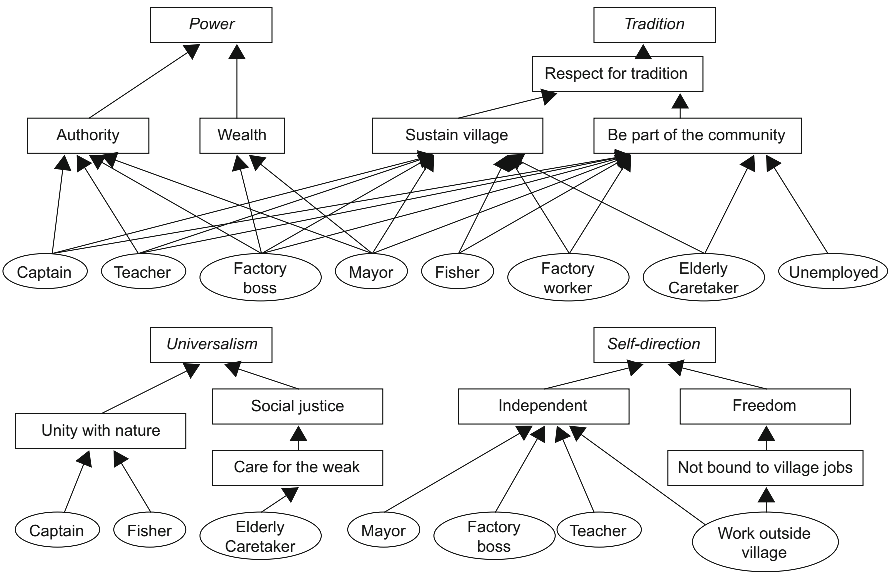


Fig. 19.3 Value tree of getting a job

picking up the possible actions. For example, if the agent wants to satisfy his $Universalism$ by doing related actions and picks being $Captain$, he only satisfies his $Universalism$ value (increasing the water level of $Universalism$ water tank) and not the $Power$ value.

19.3.2 Making Decisions

After filtering the actions by values, we have a list of actions that are value consistent. Any of these actions that get chosen by the agent comply with his value system. Among all the value-consistent action, the agent needs to pick an action that can be done at the moment. Therefore, other filters can be applied. These filters can be motivations, social norms, goals, plans, etc. The number of filters and how those filters filter down the value-consistent action set is the modeler choice. But, the result of this function is making decision about which actions have to be performed at the moment.

19.4 Validating Value Framework

In this chapter, we validate and discuss the proposed value framework, how values play an important role in human decision-making, and how decisions of individual people in a society change the overall behavior of the society. We use an agent-based model of a fishery village and show two scenarios with different abstract value settings. As it comes from the field of exploring personal values, the whole study and therefore proposing a framework for it is a qualitative study. In [9], validating a qualitative model is defined as the ability of the model to replicate the relations between variables. For instance, if the $Universalism$ value gets promoted in a society ($\tau(Universalism)$ is high), the probability of hurting the environment decreases accordingly. As described in our previous study [7], one point that we want to include in our experiments is to consider the feedback loop between society, environment, and economy. Therefore, we develop all parts and feedback between them in our simulations. The attributes and mechanics of the simulation are denoted in Table 19.1.

19.4.1 Abstract Values Implementation

There are three main action sets that use the value framework; these are job selection, event organizing/attending, and donation/not donating. We developed value trees for those actions and for the values $Power$, $Self-direction$, $Universalism$

Table 19.1 General simulation components

Age	There are four different age categories: children under 18, adults 18–64, elderly 65-74, and eldest 75 and older
Status	Adults have a status that reflects their employment the set is {*unemployed, captain, fisher, factory worker, factory boss, teacher, caretaker, worker outside, mayor*}
Ticks	There are 4 ticks per month, which makes a total of 48 ticks per year
Buildings	The buildings in the village are {*houses, school, council, factory, social care, elderly care, event hall*}. Outside the village there is another *school* and a company where agents with the status *worker outside* work
Work	Every month agents *adults* pick a job according to the value they want to satisfy. The value watertank level is increased when they keep the same job and when they switch their job
Event	Every tick *adults* and *elderly* can organize or attend an event. The organizing agents can choose between a free event (costs money) and a commercial events (generates money)
Donate	Every tick *adults* and *elderly* can choose if they want to donate to the council or not
Council	The council gets money from tax and donations and distributes it among the school, social care, elderly care and factory
Migration	Agents migrate when they are homeless and they are not happy (i.e., half of the values or more are below the threshold). A higher self-direction value then gives a higher probability of migrating

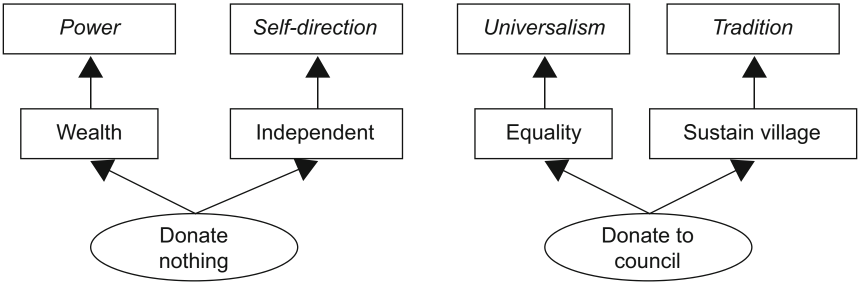
and $Tradition$. The job selection value tree is shown in Fig. 19.3. Here we see that some jobs are capable of satisfying many values like a mayor ($Tradition$, $Power$ and $Self - direction$), while other jobs have only one connected value, e.g., unemployed or factory worker (both $Tradition$). The value increase multiplier of job picking is $\sigma = 1$.

The event trees are denoted in Fig. 19.4b and show four possible actions. Organizing an event has a value increase multiplier of $\sigma = 2$ as only a small number of agents can organize an event (the maximum of events is 1 per 11 residents). Attending an event has a lower value increase multiplier; it is $\sigma = 0.2$.

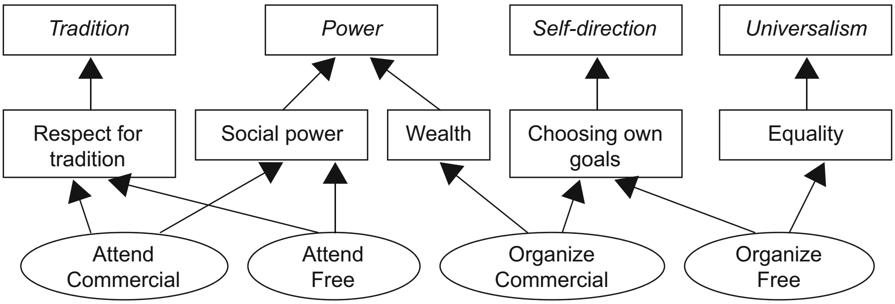
The donation trees are shown in Fig. 19.4a; there are only two possible actions here. The value increase multiplier is $\sigma = 0.2$, which is also low since donations actions can be done every tick (which is more frequent than job picking at every 4 ticks).

19.4.2 Results

We consider four values out of ten Schwartz values: $Tradition$, $Universalism$, $Self - direction$, and $Power$. These four values have been chosen because we can show that our framework works to show the dynamic relation between values. For example, Fig. 19.5 shows the dynamic behavior of the systems in two different



(a) Value tree of donation

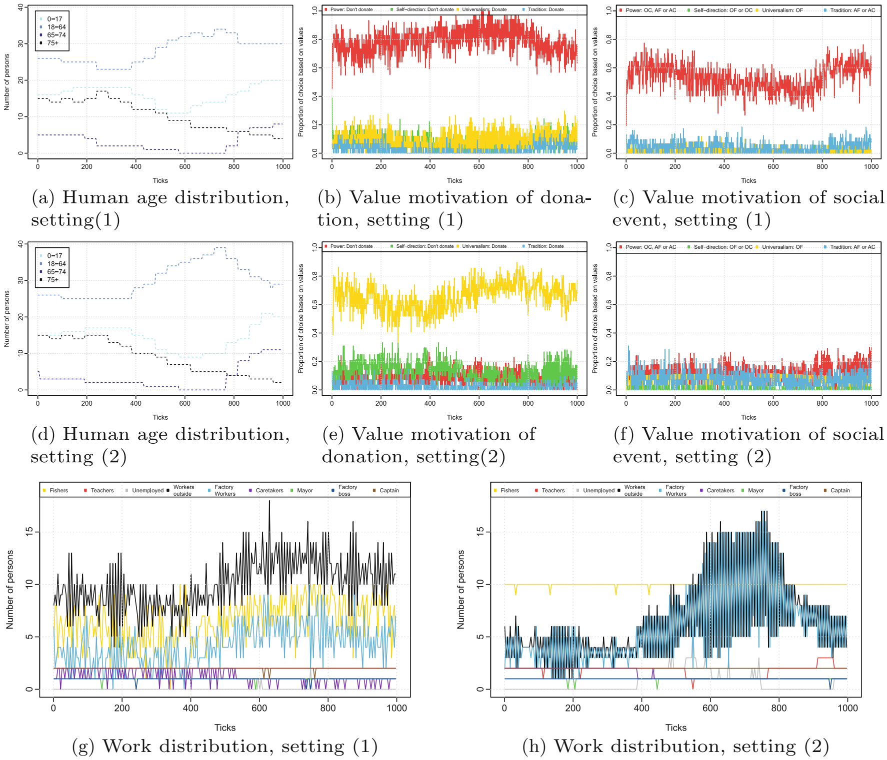


(b) Value tree of social events

Fig. 19.4 (**a**) Value tree of donation. (**b**) Value tree of social events

settings. Figure 19.5a–c, g show the system output with setting (1) when there is a high priority of $Power$ ($\tau(Power) = 90$, $\tau(Self-direction) = 50$, $\tau(Universalism) = 10$, and $\tau(Tradition) = 50$). Figure 19.5d–f, h show the behavior of the system when $Universalism$ is promoted ($\tau(Power) = 10$, $\tau(Self-direction) = 50$, $\tau(Universalism) = 90$, and $\tau(Tradition) = 50$). Having high priority for $Universalism$ means that agents need to do actions that satisfy $Universalism$ more. As shown in Fig. 19.5e, agents satisfy their $Universalism$ through donating public benefits, and there is almost always a maximum amount of fishers and captain, since this also satisfies $Universalism$. As agents have low priority for $Power$ (they do not need to satisfy $Power$ value very often), they organize commercial events and attend free events which are enough to keep them satisfied of $Power$ value.

As shown in Fig. 19.5h, most of the agents make money as they have job. So, a lot of them earn enough money that makes them capable of donating. Therefore, they satisfy their $Universalism$ value by donating in public benefits, working as a fisher, a captain, or elderly caretaker. As $Power$ importance is low (as it is the opposite value in the Schwartz value circle), the other two values need to be satisfied with the same frequency; $Self-direction$ and $Tradition$. These two values



(a) Human age distribution, setting(1)

(b) Value motivation of donation, setting (1)

(c) Value motivation of social event, setting (1)

(d) Human age distribution, setting (2)

(e) Value motivation of donation, setting(2)

(f) Value motivation of social event, setting (2)

(g) Work distribution, setting (1)

(h) Work distribution, setting (2)

Fig. 19.5 Results of simulation with two settings: (1) $p = 90, s = 50, u = 10, t = 50$ and (2) $p = 10, s = 50, u = 90, t = 50$. (**a**) Human age distribution, setting (1). (**b**) Value motivation of donation, setting (1). (**c**) Value motivation of social event, setting (1). (**d**) Human age distribution, setting (2). (**e**) Value motivation of donation, setting (2). (**f**) Value motivation of social event, setting (2). (**g**) Work distribution, setting (1). (**h**) Work distribution, setting (2)

can be mostly satisfied with picking relating jobs. Working in the factory inside the village and working in the company outside the village satisfy $Tradition$ and $Self-direction$, respectively. That is the reason we can see a fluctuation between workers in the company outside of the village and workers in the factory inside the village. There is a balance since both have free vacancies. The company outside of the village has no limit, and the factory can have a high number of employees since there is a high amount of fish coming in, because of the maximum amount of fishers.

One of the interesting simulation results is when a society is more into the $Power$ value. As we can see in Fig. 19.5g, the number of employees for the jobs factory worker, fisher, and work outside the company fluctuates but follow a general trend. The amount of fishers is lower since people hardly ever need to satisfy universalism so they mainly become fisher because of tradition. The difference in average amount of factory workers and workers in the company outside is caused by a lower amount of vacancies as factory worker because there is less fish being caught. People satisfy their *Power* value by organizing commercial events and having well-paid jobs. In this case the maximum possible number of commercial events happens all the time. So, there are more chances to attend events for villagers to satisfy their *Tradition* value by attending the events. Besides, people tend to keep their paid job as they can make enough money to cover their living cost. The importance of *Universalism* (as the opposite value of *Power* in the Schwartz value circle) is low, and there is no need to put more effort than donating in public benefit to satisfy this value. Therefore, people who do not have chance of finding a job inside the village will look for a job outside. This justifies the higher number of people who work in the company outside. The simulation code is accessible via GitHub [16].

19.5 Discussion and Future Work

Different factors impact human behavior such as values, social norms, and environmental and economic factors. However, introduced models to study human behavior rarely considered social, environmental, and economic factors altogether. Many factors are involved to capture human behavior including personal, social, environmental, and economic factors. Values are strongly connected to behavioral choices of people among personal factors. One of the well-known theories in personal values is the theory introduced by Schwartz, and it has been used by many researchers. Schwartz came up with ten general values by studying people all around the globe. Though, using Schwartz values necessitates interpreting the abstract values to concrete values related to the case study. To the best of our knowledge, there is no standard way of using Schwartz values and transform them from general to concrete values. As of yet, researchers used them and translated them according to their taste. We introduced a framework of personal values that can be used as a guideline for those who consider values to study, model, implement, and reuse previous efforts regarding values. Using the introduced framework, it is possible to model heterogeneous agents in terms of their personality and deliberation and consider various statuses consciously. For example, two different people can do the same action for different reasons, or they can react differently in the same (social, environmental, economic) situation. In our framework, we make a value tree for each value in Schwartz value theory. The root of the tree is a general value, and value gets more and more concrete till the leaves of the tree are the most concrete values that are directly linked to implementable actions. A possible actions set is assigned to each concrete value. The result of doing one of the actions in the action set is satisfying the assigned value.

In the framework, there is a relation between Schwartz values that play an important role in decision-making. Such a relation is used to capture the circular relation of Schwartz values. The framework contains making decisions according to personal values. To make a decision at each time and determine which value is more important, we used the water tank model. We assigned a water tank to each value which drains in each time step and fills whenever the assigned value is satisfied. Using such a model, agents need to satisfy all the values during the simulation time. By changing the amount of draining and filling and therefore changing the satisfaction frequency of values, we can capture different personalities. We illustrate the use of the framework by using it to build a normative architecture for developing a socio-ecological complex system. The normative architecture is a modular one that proposes developing flexible socio-ecological complex models. This architecture includes social, environmental, and economic factors, as well as decision-making process of agents. Therefore, it is possible to make a model both for micro- and macro-analysis depending on the decision of the modeler. Another aspect of this is that manipulating different factors is possible. A model may include any of the social, ecological, and economic factors. As an example of social factors, a model might contain personal values, social norms, motives, social practices, etc.

References

1. Schwartz, S.H.: In: Zanna, M.P. (ed.) Universals in the Content and Structure of Values: Theoretical Advances and Empirical Tests in 20 Countries. Advances in Experimental Social Psychology, vol. 25, pp. 1–65. Academic (1992). [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0065260108602816
2. van der Weide, T.: Arguing to motivate decisions. Ph.D. dissertation, Utrecht University (2011)
3. Dechesne, F., Di Tosto, G., Dignum, V., Dignum, F.: No smoking here: values, norms and culture in multi-agent systems. Artif. Intell. Law **21**(1), 79–107 (2013)
4. Bench-Capon, T.J.M., Dunne, P.E.: Argumentation in artificial intelligence. Artif. Intell. **171**(10–15), 619–641 (2007). [Online]. Available: https://doi.org/10.1016/j.artint.2007.05.001
5. Groeneveld, J., Müller, B., Buchmann, C., 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.: Theoretical foundations of human decision-making in agent-based land use models – a review. Environ. Model. Softw. **87**, 39–48 (2017). [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364815216308684
6. O'Brien, K.L., Wolf, J.: A values-based approach to vulnerability and adaptation to climate change. Wiley Interdiscip. Rev. Clim. Chang. **1**, 232–242 (2010).
7. Heidari, S., Dignum, F.: Value based agents for social simulation of fishery management. Social Simulation Conference (2017)
8. Bench-capon, T.J.M., Science, C.: persuasion, practical reasoning, persuasion in practical argument using value-based argumentation frameworks, J. Log. Comput. **13**(3), 429–448 (2003)
9. Mercuur, R.: Interventions on Contextualized Decision Making: An Agent-Based Simulation Study. Master thesis, Utrecht University (2015)
10. Schwartz, S.H.: An overview of the Schwartz theory of basic values. Online Read. Psychol. Cult. **2**, 1–20 (2012)
11. Tosto, D., Dignum, F.: Simulating Social Behaviour Implementing Agents Endowed with Values and Drives. Multi-Agent-Based Simulation XIII **7838**, 1–20 (2012)
12. Schwartz, S.H.: A theory of cultural values and some implications for work. Appl. Psychol. **48**(1), 23–47 (1999). [Online]. Available: https://doi.org/10.1111/j.1464-0597.1999.tb00047.x
13. Dörner, D., Bamberg, D.: A simulation of cognitive and emotional effects of overcrowding 2006 (2014)

14. Cranefield, S., Winikoff, M., Dignum, V., Dignum, F.: No pizza for you: value-based plan selection in BDI agents. In: Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI-17, pp. 178–184 (2017). https://doi.org/10.24963/ijcai.2017/26
15. Bardi, A., Schwartz, S.H.: Values and behavior: strength and structure of relations. Personal. Soc. Psychol. Bull. **29**(10), 1207–1220 (2003)
16. Jensen, M., Heidari, S.: Social-simulation-fishery-village. https://github.com/maartenjensen/Social-simulation-fishery-village.git (2018)

Chapter 20
To Stay or Not to Stay? Artificial Sociality in GRASP World

Gert Jan Hofstede and Chutao Liu

Abstract This paper describes an agent-based model that investigates group longevity in a population in a foundational way, using theory on social relations and culture. The model is the first application of the GRASP meta-model for social agents, containing elements of Groups, Rituals, Affiliation, Status and Power. It can be considered an exercise in artificial sociality: a culture-general, content-free baseline trust model from which to engage in more specific studies. Depending on cultural settings for individualism and power distance, as well as settings for xenophobia and for the increase of trust over group life, the GRASP world model generates a variety of patterns. Number of groups ranges from one to many, composition from random to segregated and pattern genesis from rapid to many hundreds of time steps. Parallels are discussed between patterns found in GRASP world and patterns found in societies that differ on individualism, power distance and heterogeneity.

Keywords Culture · Individualism · Power distance · Uncertainty avoidance · Xenophobia · Agent-based model · GRASP · Status-power theory · Artificial sociality

G. J. Hofstede (✉)
Social Sciences, Information Technology, Wageningen University, Hollandseweg, The Netherlands

UARM, North-West University, South Africa
e-mail: gertjan.hofstede@wur.nl

C. Liu
Social Sciences, Information Technology, Wageningen University, Hollandseweg, The Netherlands
e-mail: chutao.liu@wur.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_20

20.1 Introduction

Issues of group cohesion are being studied in various fields and at various scales: mergers and acquisitions, small group dynamics, organizational behaviour, conflict management, politics, geopolitics, etc. In this paper we look for common ground between all of these domains. We aim for a bottom-up way of growing groups and modelling their composition and existence through time, using the technique of agent-based modelling. Our self-confidence for attempting such a generic model is based in empirics: the cultural differences found across the world that touch all walks of life in a society, from the family to the school, workplace and political arena [1]. However, below culture there is another layer: relational life [2]. Our agents need relational motivations. I call this relational logic in agent-based models 'artificial sociality', a term coined by Gilbert and Conte [3]. In artificial sociality, both individual motives, formal social structure and cultural 'unwritten rules' matter. Cross-cultural issues turn out to be important in many socio-technical systems, because they imply different unwritten rules of the social game [1] and therefore different norms and values. When people differ in their values, norms, beliefs and rituals, but have to collaborate in a single technical environment, or using a single formal organizational structure, misunderstandings that lead to malfunctioning systems are very likely. Real life abounds with examples about the unanticipated side effects of changes in technology or in policy. Examples in socio-ecological systems include unintended pollution or environmental degradation. In socio-technical systems, they include altered opinion dynamics in networks dominated by social media versus face-to-face interaction. For instance, social media are changing opinion dynamics, while culture also shapes social media usage [4].

We draw our agents' relational drives from Kemper's generic status-power theory [5]. Since this is the first implementation of such a model, and truth emerges more readily from error than from confusion, it is kept very simple; our hope is that others will take these ideas further. While the thinking behind GRASP is available [6], implementations are not. This is why the theory development covers a bit more ground than the implemented model; it shows potential areas for development of GRASP world. Still, we omit important areas: there are no content issues, no cognitive complexity of individuals, no cross-group status-power relations and no evolution.

As far as we know, GRASP world is the first model of its kind. It is close to abstract social models that rely on a limited set of concepts and mechanisms for modelling emergence of system-level pattern, e.g [7–9]. It differs in its reliance on existing social science theory. GRASP paints with a slightly finer brush, trying to add cultural variation to a model that is generic across all people. It does not pretend to be anything more than a first attempt. This means that no validation beyond plausibility is possible. Any real-world application will contain features that are absent from GRASP world, which could affect the dynamics.

20.1.1 To Stay or Not to Stay?

To resolve collective action issues, actors frequently have to invest effort in cooperation. The willingness of actors to cooperate is strongly related to their expectations about whether others will cooperate [10]. The absence of trust leads to lines of reasoning such as 'if I preserve this common resource, someone else will take it, so I might as well take it myself'. This could lead to a vicious circle. Partners would then be forced to either stop collaborating altogether or try to subdue others into compliance.

People do not just engage in collective action with anyone; they can choose their partners to some extent, for instance, within a fixed in-group. A second mechanism is that people can negotiate the extent of their pro-social behaviour in promises or contracts and thus voluntarily mutually expose themselves to compliance. This is what happens in a commons [11], avoiding the selfish overconsumption that results in a 'tragedy of the commons'.

How do actors know whether they have the right partner? Perhaps their most crucial option, relying on their experience, is leaving a bad partner or group. This is the main focus of the model to be discussed here. Often the question of whether to stay with a group is framed as 'trust', certainly in individualistic cultures, where changes in group membership are frequent. However, trust is a multifaceted phenomenon, and its meaning could vary across cultures [12]. GRASP world is naive about the question whether the agents themselves would think of 'trust'. Neither does it model any shareable resources. Only the agents' action counts: they leave, or they stay. How difficult this decision is, many readers will know from their own experience in jobs or personal relationships. At the level of nations, it is apparent from issues such as boundaries and fission and fusion dynamics between empires and states. As an illustration, consider the history of Europe, with EU membership being the latest in an endless succession of changes.

This first implementation of GRASP world uses the main GRASP components (*Groups, Rituals, Affiliation, Status, Power*). Agents need the following:

- A motive for gathering in various groups (*Groups, Affiliation*)
- A way of conferring status (in this case a content-free mutual exchange of status; *Ritual, Status*)
- A motive for leaving a group (being 'unhappy': receiving less status than one deems oneself worthy of; this motivates an agent to search for alternative groups that might confer more *Status*)
- A way of settling a status offence (picking a fight: *Power* contest). The outcome changes their social status: the winner gains status, the loser goes down in status (as in Hemelrijk's DomWorld; see [9, 13]).

Additionally, GRASP world contains constructs for dealing with culture and trust in the simplest ways, discussed below. These modify the strengths of the motives.

20.1.2 Groups Across Cultures

The organization of society into groups and institutions, companies and governments differs across cultures. This depends on many factors, of which in GRASP world only culture remains, since its world is undifferentiated. For a bird's eye perspective on culture, we take Hofstede's model with its six dimensions: individualism, power distance, masculinity, uncertainty avoidance, long-term orientation and indulgence. Two of these are of obvious relevance for the probability of leaving a group. These are individualism and power distance [1, 14]. Other dimensions may play a role too; of these we shall discuss uncertainty avoidance in the context of xenophobia and familiarity.

A society's individualism is a measure of the independence of individuals. It ranges from voluntarist, often volatile relationships in individualistic societies, versus life-long interdependence of group members in collectivistic ones. In the former case, leaving a group is easy; in the latter, it is almost impossible.

A society's power distance has to do with acceptance of differences in social status and power. In a hierarchical (i.e. with large power distance) culture, individuals that receive less status are likely to accept this as natural. Powerful members in societies of large power distance also feel entitled to using their power: 'might makes right'. Thus, in a culture of large power distance, those who receive less status are likely to accept this and not have a motive for leaving the group. They might wish to leave in their hearts, but not feel empowered to do so; or they might simply not consider the option. In an egalitarian (i.e. with small power distance) culture, everyone is empowered to decide for themselves.

The plausibility of these ideas is corroborated by cross-cultural comparative findings. We shall mention a few across domains. This also makes the point that GRASP world is applicable across many application domains.

Consider, e.g. Fig. 4.1 in Hofstede, Hofstede and Minkov [1] that plots countries on the dimensions of individualism against power distance. Countries with open political systems that change smoothly over time are concentrated on the side of egalitarian, individualistic cultures. Countries with long-standing regimes interspersed with revolutions are found on the side of large power distance and collectivism. Or consider the findings on consumer behaviour in de Mooij [15], who found, e.g. a strong relationship across EU countries between collectivism and brand loyalty. Trying a new brand apparently has a connotation of leaving one's group.

A more proximate phenomenon is the behaviour of tourists or reception guests in multinational settings. Those from more collectivistic cultures are more likely to be found in in-group clusters.

The fact that differences in group behaviour are culture-mediated is also borne out by studies among children showing that for them too, the probability of leaving a group varies with individualism and power distance [16, 17].

Table 20.1 Hypothesized sources of continued group membership according to culture

		Individualism	
		Individualistic	*Collectivistic*
Power distance	*Small*	Bonus schemes, certification, contract, empathy, legal enforcement, monitoring, perceived match, price premiums, social norms, transparency	Identification Kinship Obligation Perceived match Reputation Social norms
	Large	Dependence, hierarchy, honour, sense of duty, social norms	Allegiance, hierarchy, kinship, reputation, sense of duty, social norms

20.1.3 Trust Across Cultures: Individualism and Power Distance

In collectivistic cultural settings, trust in formal, impersonal institutions tends to be low [1]. Instead, sources of trust are found in the relationship in the in-group or in informal institutions such as norms. If the societal context is also hierarchical, implicit hierarchy will determine loyalties.

In individualistic societies, everybody is a priori supposed to be equally trustworthy. Trust can be based on incentive control or on positive emotions. Impartial third parties and institutions will be accepted, especially if the society is also egalitarian. If the society is also hierarchical, a sense of filial duty could play a role.

A cross-European study on trust placed in sellers in food supply networks [12] yielded a typology that is useful for our purposes. The typology was inspired by Nooteboom [18]. It opens the way to a discussion of relationships between trust and the culture-bound organization of society. Table 20.1 is based on it, with additions. It shows that many of the contemporary trust mechanisms such as contracts fit an individualistic, egalitarian culture that describes the prototypical Western, highly institutionalized society. To the extent that the society is less egalitarian and more collectivistic, other forms of trust enforcement will be important in practice, even if they are not codified in laws. Social norms exist everywhere, only not the same ones.

Table 20.1 uses term charged with content matter. In our GRASP world, there is no content. Therefore, we ‘zoom out’, leaving only the most basic feeling of ‘happiness’. This is meant in Kemper’s sense as ‘having received appropriate status’. The various sources of trust from Table 20.1 are collapsed onto agent ‘happiness’. Table 20.2 shows the influence of individualism and power distance on leaving a group in GRASP world.

Table 20.2 Reasons for leaving a group depending on individualism and power distance. A ‘big shot’ is the group member with the most social status, if above one’s own

		Individualism	
		Individualistic	*Collectivistic*
Power distance	*Small*	Leave if unhappy	Leave if very unhappy
	Large	Leave if unhappy and no ‘big shot’ present	Do not leave

20.1.4 Social Identity, Xenophobia and Familiarity, and Uncertainty Avoidance

Trust is related to social identity and shared practices [19]. Predictable behaviour by others and trust in them go hand in hand. This is observed even in individualistic societies or in mixed-culture settings: one tends to trust people of the same social identity, regardless of the culture that they might have. The difference is that social identity is visible on the outside, while culture is not.

The dimension of culture Uncertainty Avoidance is relevant here. An uncertainty avoiding culture is one in which people feel anxious in situations that they are unfamiliar with. This has the effect of making them slow to adopt unpredictable changes, such as accepting members from unfamiliar groups, and more rigid in their rituals. Strong uncertainty avoidance in a society is associated with distrust of strangers and love of familiar ritual. For the dynamics of trust in a group, the above means that people could be expected to become less likely to leave a group the longer they have already been in it. This effect should be stronger in more uncertainty avoiding societies, where familiarity is more strongly reassuring [1] (Chap. 6). Still, these two trust-related aspects of group life (xenophobia and familiarity) are conceptually distinct. We shall take them both on board separately.

To model xenophobia we introduce a simple scalar parameter that denotes different practices or social identities, visible to outsiders. In the real world, one could think of hairstyle, colour, clothing, greetings, food habits and so forth. We call it ‘norm meme’ for ‘Nexus of Relational Memes’. The acronym ‘norm’ could be confusing, and neither ‘social identity’ nor ‘practice’ fit the bill, so we refer to it as ‘shade’, since agents’ shade is its visible form in the interface. In any of the quadrants of Table 20.2, xenophobia could lead to two potentially opposing trends. If we have a homogeneous society, uncertainty avoidance could be associated with smoothly oiled operations in which all take pride in being predictable. If, on the other hand, we have a heterogeneous society with varying practices and social identities, xenophobia could be associated with segregation along the lines of differences in behaviour or in visible group affiliation.

Familiarity depends on experience in a group. Agents only interact within their group, until they leave it and join a random other agent or group. Agents in a group acquire the same ‘aura’ colour. Depending on the weight of the parameter ‘dyadic-trust-factor’, groups can generate familiarity-based trust, making their members progressively less likely to leave.

20.2 The Model: GRASP World

20.2.1 Description

For reasons of space, the following more technical description is incomplete; more details can be obtained from the authors. GRASP world is built with NetLogo (Wilensky 1999, version 6.0.2) and is available on COMSES [20], including an ODD + D description [21]. Model interface is shown in Fig. 20.1. Space limitations prohibit showing the details in this chapter.

Agents with different shades randomly scatter on an undifferentiated ground. One tick represents one time of interaction, and simulation is run for 1000 ticks. The length of a tick is deliberately undefined. It could be a second, and the simulation could show a schoolyard or a reception. It could be a year, and the simulation could represent a political landscape or a social network. For each tick, each individual goes through the process ‘Play with group mate’, during which an individual and his/her communication partner for this tick confer status to each other and perceive status affront. This mutual status-conferral and perception process affects the pair’s emotions (‘happy’, ‘neutral’ or ‘angry’) and consequently influences the group dynamics.

The status conferral for each is determined by picking a random number from a normal distribution with a mean of 0.5 and a standard deviation set by ‘status-conferral-stdev’. Perceived status affront is calculated as follows (‘my’ and ‘mate’ symmetrical):

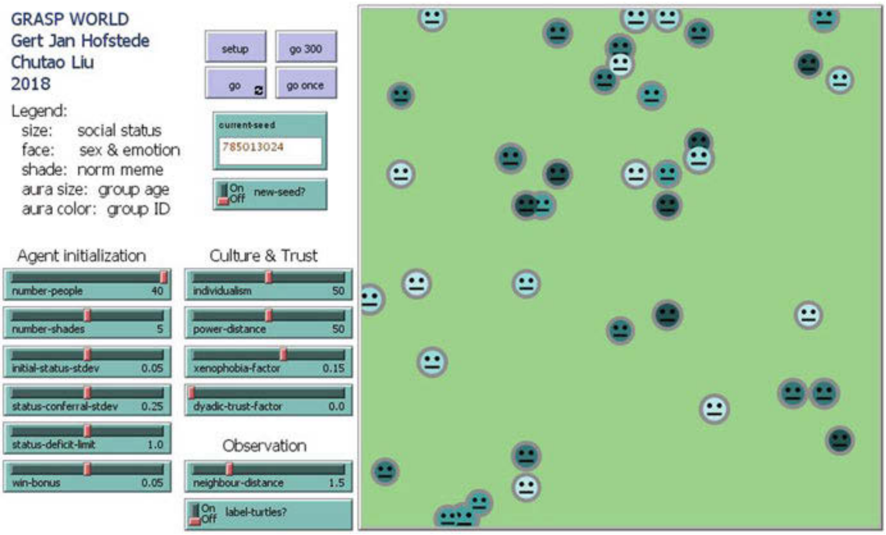


Fig. 20.1 Model interface at $t = 0$. All agents have equal status indicated by avatar size, and there are five shades

$$My - perceived - status - affront$$
$$= \text{(my-status-conferral–mate-status-conferral)}$$
$$+ \text{norm -misunderstanding–submission–dyadic-trust.}$$

where

Norm-misunderstanding = *xenophobia-factor* * | *my-shade – mate-shade* |
Submission = power-distance/100 * (mate-status – my-status)
Dyadic-trust = (dyadic-trust-factor/10) * time-in-group-together

For the ‘decide to leave or not’ process, the group member with highest status (‘bigshot’) will decide first. The decision mechanism is the same for everyone and is based on individualism, person’s own emotion and other group members’ emotions. If the bigshot decides to leave the group, then other ‘ordinary’ members in the group can make their decisions and leave as they want. If the bigshot decides to stay, then other members have to overcome their submission to the bigshot to be entitled to leave. The group will be dismissed if less than two members remain.

20.2.2 Model Analysis

Sensitivity analysis was performed using one-factor-at-a-time method (OFAT) on all parameters, with non-varied variables at default value. Due to limited space, we present only the two most sensitive ones. Number of groups and status distribution (standard deviation) among agents are selected as outcomes. Their sensitivity to xenophobia-factor (with a population consisting of 5 norm shades) and dyadic-trust-factor are presented in Figs. 20.2 and 20.3, respectively.

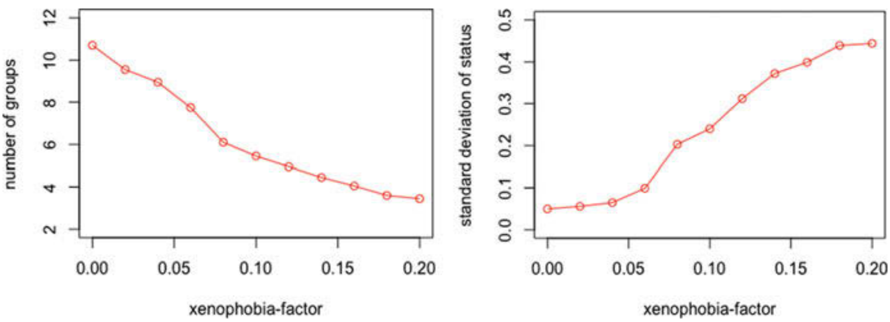


Fig. 20.2 The sensitivity of ‘number of groups’ (left) and ‘standard deviation of status among agents’ (right) to xenophobia-factor. Each data point is the average value of 20 repetition runs of 1000 ticks at corresponding parameter settings

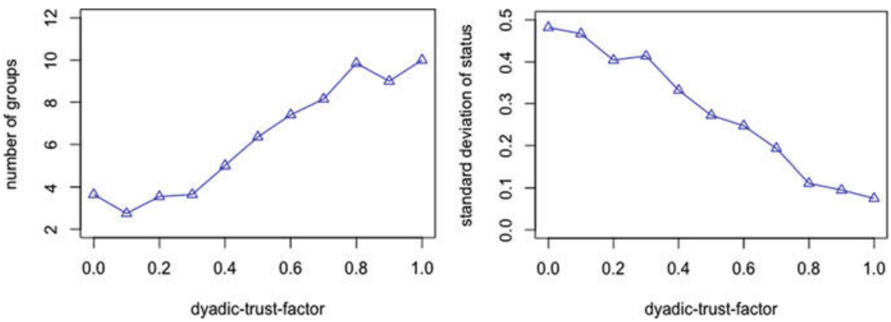


Fig. 20.3 The sensitivity of 'number of groups' (left) and 'standard deviation of status among agents' (right) to dyadic-trust-factor. Each data point is the average value of 20 repetition runs of 1000 ticks at corresponding parameter settings

Table 20.3 Results across culture in the absence of xenophobia

		Individualism	
		Individualistic	*Collectivistic*
Power distance	*Small*	One happy group	Many stable happy groups
	Large	One group with status fights that can break up ultimately after >1000 ticks	Many stable groups with infighting

Both outcomes are sensitive to xenophobia-factor and dyadic-trust-factor. The increase in xenophobia leads to fewer groups and more obvious status differentiation. Since xenophobia positively contributes to agents' perceived status affront, higher xenophobia-factor will cause more fights and consequently lower group stability. Under frequent fights, winners are able to accumulate high status, while losers tend to be suppressed over time, which intensifies status differentiation. In contrast, higher dyadic-trust-factor reduces the possibility of fights and thus stabilizes groups. When dyadic trust is high, nearly no fights break out and agents seldom gain or lose status during their interactions with each other. Hence, the status differentiation and model dynamics in general are less with higher dyadic-trust-factor.

20.2.3 Model Results

Table 20.3 gives the trends observed when only considering individualism and power distance. Since tick 1 always yields a large number of small groups, which may or may not evolve subsequently, Table 20.3 shows that ability to leave boosts evolution of groups. Xenophobia leads to more fighting if there are different shades, while in combination with dyadic-trust, it can still yield stable, usually shade-segregated groups.

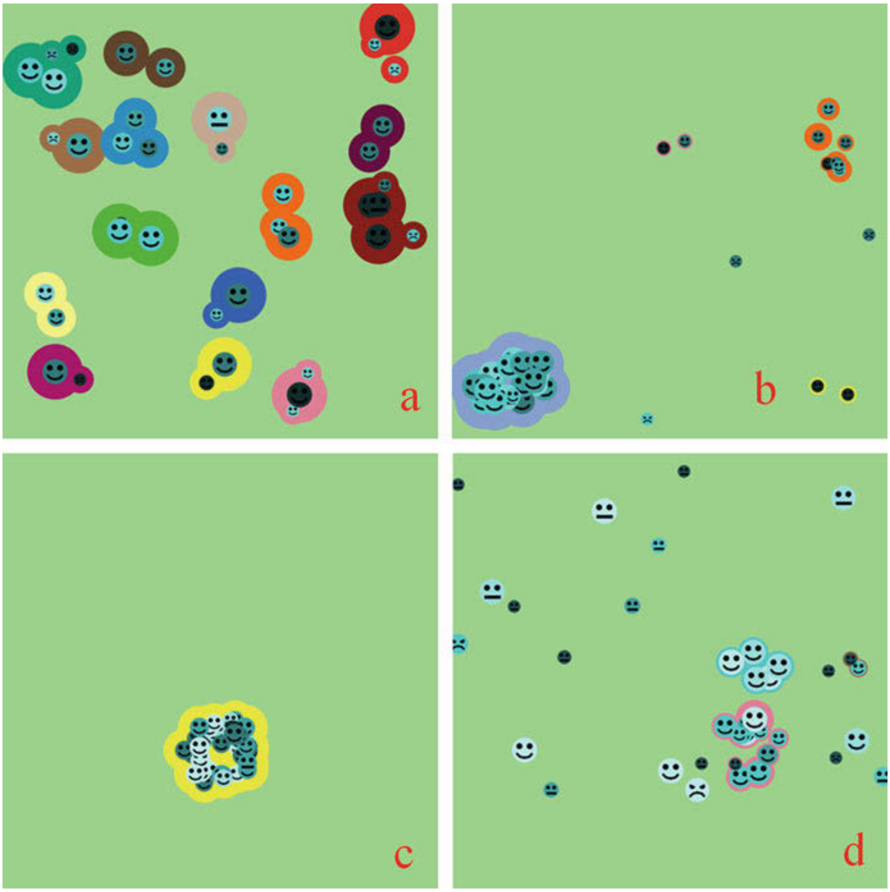


Fig. 20.4 (**a**–**d**) Group dynamics at $t = 1000$ under different parameter settings. The aura around the individuals marks their group. Individuals with same aura colour are in the same group. The longer an agent stays in a group, the larger its aura size grows, until 20 ticks. Individuals with no auras are alone at the moment

Depending on the settings of trust-factor, xenophobia-factor and the two cultural factors and the shades, a number of patterns can occur. Some are stable; some show periodic upheaval. We show five of these patterns in Figs. 20.4a–d and 20.5. All non-mentioned parameters are at their defaults.

- *Pattern 6a*: low individualism (10), large power distance (90), no xenophobia, no dyadic-trust
- *Pattern 6b*: high individualism (90), large power distance (90), no xenophobia, no dyadic-trust
- *Pattern 6c*: high individualism (90), medium power distance (50), no xenophobia, no dyadic-trust

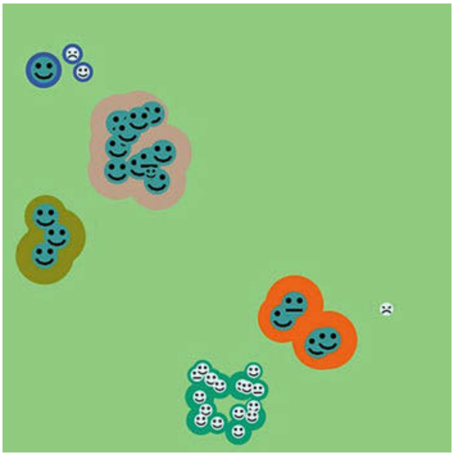


Fig. 20.5 Pattern e, Happy Segregation. The picture shows part of the interface at $t = 600$. This population with two widely different norm shade subpopulations self-organizes into a segregated system, with usually two groups per shade, agents staying in a group for more than 40 ticks, little fighting, but occasional breakdowns

- *Pattern 6d*: high individualism (90), small power distance (10), medium xenophobia (0.2), low dyadic-trust-factor (0.1)
- *Pattern 7e*: high individualism (80), medium power distance (50), medium xenophobia (0.2), medium dyadic-trust (0.3)

Pattern a: Many long-standing groups Under low individualism, agents are likely to stay in groups. Large power distance further contributes to group stability by keeping low-status members in. Hence, many small groups created in tick 1 sustain themselves over time. Status differentiation occurs among agents within groups, due to fights.

Pattern b: Stable elite groups, volatile others Under high individualism, agents tend to leave the group easily. Even large power distance is not strong enough to stop them. Fewer but larger groups are formed, with clear status differentiation between them. If there is medium xenophobia, agents of high status tend to form all-elite groups with similar shade. These groups tend to be longer-lived than others, since other groups can be 'blown up' when they are colonized by an escapee from an elite group.

Pattern c: One happy group Without xenophobia, and under small power distance, highly individualistic agents generally form one single group, even when dyadic trust = 0. Such a large group is stable over time and inclusive regarding different shades. There are hardly any fights. As a result, status differentiation occurs slowly, and the majority of group members are happy.

Pattern d: Nothing lasts for long This can occur when xenophobia is increased in parameter settings that show pattern c, regardless of power distance. High individualism plus small power distance makes it easy for agents to leave the group, so many agents are alone. The lack of trust also leads to less cohesive groups that do not last long.

Pattern e: Happy segregation Figure 20.5 highlights one more pattern across time. Hierarchy establishes within 100 ticks, although not yet linked to shade. Gradually, fighting and group splitting has the side effect of coupling status with shade. Then, the system stabilizes as a 'class society'. The system remains metastable. Runs with the same settings but more shades often either default to pattern c or d.

Depending on settings and number of shades, other patterns obtain, e.g. smaller single-shade groups circling around large melting-pot groups, and various dynamics occur.

20.3 Discussion and Conclusion

GRASP world and culture Recognizable group life patterns are generated by a rather simple model based on foundational social scientific theory. They bring to mind societal patterns that are particularly apparent to people with cross-cultural experience.

Individualism: Perhaps counter-intuitively, individualistic agents find it easier to unite, because they can leave a prior group in which they received bad treatment. There are phases of chaos in a simulation run in which every agent moves around from one to the other. This allows the agents to self-organize until finding acceptable configurations. These could be status- or shade-sorted, depending on settings. There could be a parallel in the spontaneous forming and dissolving of all kinds of groups in individualistic societies, with usually limited lifetime.

Power Distance: The long-term patterns generated by status formation, in which initially stable groups become undermined, are interesting. If, in a real-world analogy, the status hierarchies were destroyed during such upheaval, this pattern could repeat itself, in analogy to revolutions and finite periods of stability following them. This is indeed what is argued by Turchin [22].

GRASP world, social identity and trust The number of 'norm meme' parameter values indicated by shade and perceived through xenophobia is sufficient to modify the base patterns from Fig. 20.4 considerably. This yields patterns a through e. Similar patterns can be recognized in group dynamics in organizations and societies. Experimenting with scale of norm difference and number of groups could lead to other plausible patterns.

The familiarity parameter for trust emanating from a group's age seems less interesting, since it just tends to freeze any pattern that occurs.

Interpreting the patterns A theory-based, content-free virtual world such as GRASP world cannot reproduce any specifics of any situation. Yet it can show the recurrent patterns in societies that are captured by Blaise Pascal's famous phrase 'Plus ça change, plus ça reste la même chose' (the more things change, the more they stay the same). Let us interpret the five patterns with some freedom of imagination. The various proximate mechanisms and value-laden words used in these interpretations are not implemented in GRASP world; but the idea is that they are all manifestations of status-power dynamics under cultural variation, and those two actually form the heart of GRASP world. So from a mental helicopter, we can recognize our real world in GRASP world.

Pattern a: many long-standing groups is characteristic of most societies in the world. Families or clans are long-lived almost everywhere. Only in Western societies, and more so in higher than in lower socio-economic classes, is there a lot of mobility between groups. If there are differences in social identities between subgroups, there will either be segregation, or the differences will be strongly suppressed.

Pattern b: stable elite, volatile others characterizes many societies with affluent elites and a large, rather impoverished, working class. The elite form stable networks and 'pass the ball to one another'. Temporary contracts force the working class into a volatile, often migratory existence, although they might, if given the chance, prefer to stay in pattern a.

Pattern c: one happy group describes societies in good times. Here we find the 'Global Village' metaphor enacted. This is particularly likely to happen in individualistic societies. For instance, at a major sports victory, or when a king is crowned, Scandinavian societies and the Netherlands feel like one happy family. Individualistic societies can also join in bad times and be 'happy' in terms of giving one another positive feeling or other forms of support.

Pattern d: nothing lasts for long characterizes individualistic societies in the absence of a reason for group solidarity. There is a lack of loyalty. Opportunism leads to volatile marriages, job-hopping and other forms of cross-group mobility in search of social status. Typical Western business practices of top-level executives or venture capitalists can show this pattern.

Pattern e: happy segregation has occurred and still occurs across religious or ethnic factions that coexist in the same society, but have inhibitions when it comes to mixing. In the Netherlands, these were called pillars, '*zuilen*'. The two sexes could also coexist in this way to some degree.

We conclude that if a simulation requires human population with credible group dynamics over time, GRASP world citizens, given appropriate parameters, could fit the bill.

Further work Among myriad possible enhancements, these seem promising to us:

1. The model itself can be enhanced, e.g. with individual differences between agents, or with multicultural populations. Notably, differences in status worthiness (attractiveness) of agents, possibly including heterogeneity in perception ('beauty is in the eye of the beholder'), could alter the dynamics; or agents could differ in cultural traits within one run.

2. The relationship between dimension of culture and model concepts can be further investigated. This includes thinking how other dimensions of culture could play a role [23]. However, adding model complicatedness exacerbates validation difficulties.
3. Cross-group status-power dynamics can be added. Here, dynamics between groups similar to those currently implemented between individuals could obtain; groups would also have a sense of social status worthiness and be motivated to use power if they felt they received too little status.
4. We could give the agents something to do. In other words, the model's mechanisms can be re-used for more targeted studies in more realistic simulated worlds that include context. This would yield an avalanche of richer concepts, for instance, on the mechanisms in Table 20.1, where a closer fit between model concepts and empirical attributes of the situation could be attempted. One such attempt has been made by giving the agents a common pool to govern. They have norms over how to use the resource. This shows intriguing emergent patterns.

The agents could reproduce, and culture could be modified across generations, based on a mix of parental and societal influence. This would open up a whole new area of research into culture's causes [24].

Conclusion It seems promising to develop GRASP world, as a tool for furthering the development of artificial sociality. It can throw light on societal dynamics and on group dynamics within societies. GRASP world agents could also be useful for researchers needing to populate their virtual world with simple, socially plausible virtual agents.

Acknowledgements The paper benefited from the comments of anonymous reviewers.

References

1. Hofstede, G., Hofstede, G.J., Minkov, M.: Cultures and Organizations, Software of the Mind. McGraw-Hill (2010)
2. Hofstede, G.J.: Culture's causes: the next challenge (distinguished scholar essay). Cross Cult. Manag. Int. J. **22**(4), 545–569 (2015)
3. Gilbert, N., Conte, R.: Artificial Societies: The Computer Simulation of Social Life. University College London Press, London (1995)
4. Choi, K.S., Im, I., Hofstede, G.J.: A cross-cultural comparative analysis of small group collaboration using mobile twitter. In: Computers in Human Behavior, vol. 65, pp. 308–318, 2016/12/01/, (2016)
5. Kemper, T.D.: Status, Power and Ritual Interaction; A Relational Reading of Durkheim, Goffman and Collins. Ashgate, Farnham (2011)
6. Hofstede, G.J.: GRASP agents: social first, intelligent later. In: AI & Society, pp. 1–9 (2017)
7. Deffuant, G., Carletti, T., Huet, S.: The leviathan model: absolute dominance, generalised distrust, small worlds and other patterns emerging from combining vanity with opinion propagation. JASSS. **16**(1), (2013)
8. Grow, A., Flache, A., Wittek, R.: An agent-based model of status construction in task focused groups. J. Artif. Soc. Soc. Simul. **18**(2), 4 (2015)

9. Hemelrijk, C.K.: Simple reactions to nearby neighbors and complex social behavior in primates. In: Menzel, R.F., Fischer, J. (eds.) Animal Thinking: Comparative Issues in Comparative Cognition, pp. 223–238. MIT Press, Cambridge (2011)
10. Pretty, J.: Social capital and the collective management of resources. Science. **302**(5652), 1912 (2008)
11. Wilson, D.S., Ostrom, E., Cox, M.E.: Generalizing the core design principles for the efficacy of groups. J. Econ. Behav. Organ. **90**(S), 21–32 (2013)
12. Hofstede, G.J., Oosterkamp, E., Fritz, M.: Trust in the agri-food sector: a typology with a cultural perspective. In: Canavari, M., Fritz, M., Schiefer, G. (eds.) Food Supply Networks: Trust and E-business, p. 195. CABI, Wallingford (2016)
13. Hemelrijk, C.K.: The use of artificial-life models for the study of social organization. In: Thierry, B.S., Singh, M., Kaumanns, W. (eds.) Macaque Societies: A Model for the Study of Social Organization, pp. 295–313. Cambridge University Press, Cambridge (2004)
14. Minkov, M.: Cross-Cultural Analysis: the Science and Art of Comparing the World's Modern Societies and their Cultures. Sage, Los Angeles (2013)
15. de Mooij, M.: Global Marketing & Advertising; Understanding Cultural Paradoxes, 5th Expanded edn. Sage (2018)
16. Martínez-Lozano, V.S.-M., José, A., Goudena, P.P.: A cross-cultural study of observed conflict between young children. J. Cross-Cult. Psychol. **42**(6), 895–907 (2011)
17. Hofstede, G.J., Student, J., Kramer, M.R.: The status–power arena: a comprehensive agent-based model of social status dynamics and gender in groups of children. AI & Soc. 1–21 (2018)
18. Nooteboom, B.: Trust: Forms, Foundations, Functions, Failures and Figures. Edward Elgar Publishing, Cheltenham (2002)
19. Brown, R.: Group Processes: Dynamics Within and Between Groups, 2nd edn. Blackwell, Malden (2000)
20. Hofstede, G.J.: "GRASP world" (Version 1.0.0). CoMSES Computational Model Library. Retrieved from: https://www.comses.net/codebases/f1089671-38da-4cca-88c5-711e936b2ada/releases/1.0.0/ (2019, April 16)
21. Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., Schwarz, N.: Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. Environ. Model Softw. **48**, 37–48 (2013)
22. Turchin, P.: Historical Dynamics: Why States Rise and Fall. Princeton (2003)
23. Hofstede, G.J.: Theory in social simulation: status-power theory, national culture and emergence of the glass ceiling. In: Social Coordination: Principles, Artefacts and Theories, pp. 21–28, Exeter (2013)
24. Hofstede, G.J.: Culture's causes: the next challenge. Cross Cult. Manag. Int. J. **22**(4), 545–569 (2015)

Chapter 21
Simulating a Direct Energy Market: Products, Performance, and Social Influence

Sascha Holzhauer, Friedrich Krebs, and Christoph Nölle

Abstract We present modelling results of a direct, high-frequency energy market (DEX) where many heterogeneous (individually configured), situated (distributed), generation and consumption units of various types (photovoltaic systems, combined heat and power plants, heat pumps, wind energy converters, etc.) are able to trade energy and react to fluctuations in real time. These prosumers utilise so-called energy management gateways (EMG) which automatically interact with the DEX, following bidding strategies and preferences chosen by their owners. We simulate the market performance with an agent-based model representing prosumers as heterogeneous, autonomous entities. Results indicate that the direct market is well able to reflect scarcity through prices promptly, and sets of market products can be defined to suit participants' needs.

Keywords Energy system · Energy market · Bidding strategies · Prosumer · Agent-based modelling · Demand side management

21.1 Introduction

The future energy system is characterised by a substantial increase of distributed power generation by renewables [1], meaning a shift of the system's core characteristics: Whereas in the past, energy was generated by relatively few power plants as demanded and distributed through a hierarchical and unidirectional grid, nowadays more and more distributed PV plants and wind turbines generate fluctuating energy as they depend on meteorological conditions. Households become "prosumers" as they not only consume energy but also generate, e.g. by PV plants. At the same

S. Holzhauer (✉) · F. Krebs
Research Group Integrated Energy Systems, University Kassel, Kassel, Germany
e-mail: sascha.holzhauer@uni-kassel.de; fkrebs@uni-kassel.de

C. Nölle
Fraunhofer Institute for Energy Economics and Energy System Technology, Kassel, Germany
e-mail: christoph.noelle@iee.fraunhofer.de

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_21

time, the demand for electricity rises as the mobility and heat sector are going to be decarbonised. Besides investments in renewable power generation and storage, increasing demand and fluctuating generation require demand side management, i.e. means to align energy consumption to energy generation.

Energy prices may indicate surplus and scarcity and incentivise according energy consumption. However, due to access restrictions, the current energy market is characterised by few participants dealing large volumes of energy for relatively long intervals. Therefore, most consumers are cut off from price fluctuations and unable to react to volatile generation in real time. The integration of prosumer entities into future electricity markets requires a shift from centralised control to decentralised self-organisation [2].

We develop a software architecture design for the deep integration of the technical and economical processes of the future energy system. By harmonizing and standardisation, presently loosely connected software solutions in the different sectors of the energy system could be merged into a "Self organising energy automation system" (SOEASY). An experimental platform allows us to study the complex interplay of economic and technical processes of generation and transmission of electric energy in a close-to-reality simulation, utilizing agent-based modelling.

While ABM has been widely applied in the diverse subdomains of energy systems [3, 4], the research focus was either on actor behaviour in wholesale electricity markets [5, 6] or on the opposite end of the electricity supply chain, namely, electricity-related (household) behaviours [7–9]. The investigation of the market integration of decentralised generation and flexible consumption combines the two perspectives and is an emerging field of energy-related ABM studies. For instance, in [10], mobility behaviours are assigned to electric vehicles that seek to minimise charging costs by reacting to variable prices. The neighbourhood exchange of electricity to reduce electricity costs by dynamical demand adjustment has been analysed by [11].

In the paper at hand, we report early simulation results of a direct, high-frequency energy market (SOEASY-DEX) where many heterogeneous (individually configured), situated (distributed), generation and consumption units of various types (photovoltaic systems, combined heat and power plants, heat pumps, wind energy converters, etc.) are able to deal energy and react to fluctuations in real time. These prosumers utilise so-called energy management gateways (EMG) which automatically interact with the DEX, following bidding strategies and preferences chosen by their owners. We simulate the market performance with an agent-based model representing prosumers as heterogeneous, autonomous entities.

21.2 Method

Figure 21.1 gives an overview of the system's architecture. The SOEASY-DEX server allows clients to submit their asks of energy supply and demand via a

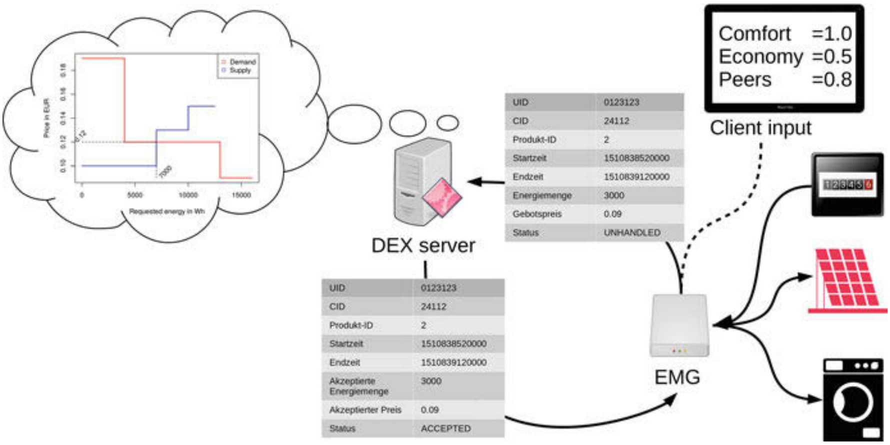


Fig. 21.1 Schematic overview of the direct market exchange (DEX) including the client-controlled, automated energy management gateway (EMG) and protocols to communicate asks

REST interface, performs uniform clearing, and provides an interface for clients to obtain results of the clearing process. It furthermore controls the real net energy consumption by processing metre readings and determines fees for untraded energy and missing meter readings.

The DEX provides a set of market products with different delivery periods and auction settings, including auction opening and closing times as well as auction intervals. Consequently, auctions may take place several times for the same delivery period, allowing clients to adjust their requests if not successful.

Clients are assigned a set of components that either generate energy (PV plant, wind turbine), consume energy (load profile), or store energy. The EMG is able to forecast consumption and generation and then buys and sells electricity at the DEX as required in an automated, optimal way. It considers provided market products and follows bidding strategies as selected by autonomous agents who represent human actors. The EMG may consider user preferences such as price sensitivity, comfort, and eco-friendliness to shift loads such as charging electric vehicles or heat pumps.

The simulation framework allows the analysis of clearing prices and traded energy as well as individual evaluations of bidding strategies and performance of scenarios that differ regarding client configuration (i.e. expansion of heat pumps) and client preferences.

21.3 Results

We simulated the market behaviour of 30 heterogeneous clients on the direct energy exchange with different market products. The products vary by the delivery interval,

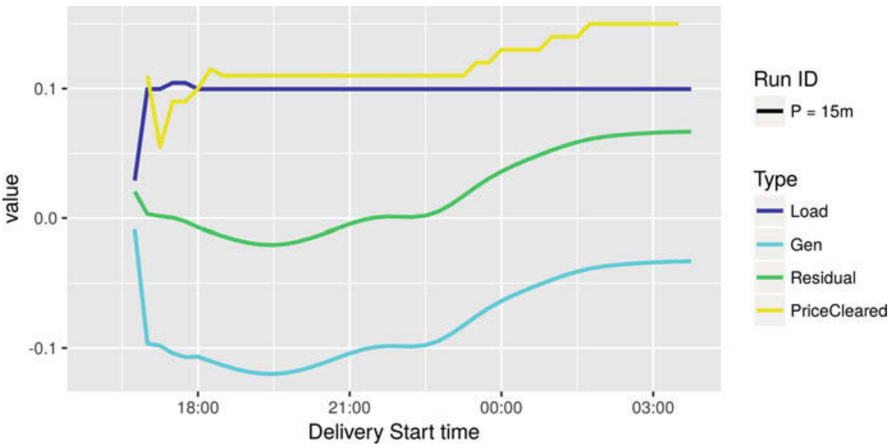


Fig. 21.2 Constant load, generation (modelled as negative load), and residual of requested energy as well as clearing price for 30 clients and a market product with delivery interval = 15 min

i.e. clients may bid a certain amount of energy generation or consumption during 5 or 15 minutes.

Whereas loads are constant for reasons of easier analyses, generation follows a time-variant profile to simulate volatile energy sources from wind turbines and PV. Both are calibrated with an annual energy figure which is randomly assigned each client.

As shown in Fig. 21.2, the resulting, increasing clearing price reflects the volatile, here decreasing electricity generation well.

The benefit of a flexible definition of market products is illustrated by Fig. 21.3. The market with two products allows participant to react to fluctuations in consumption and generation: they may forecast and ask for longer delivery products accordingly and then balance deviations in nearly real time by short-term products. With more product options, the resulting clearing price is slightly lower.

21.4 Discussion

Innovative, decentralised, high-frequency markets have a great potential to boost demand side management, but the complex nature of the energy system requires thorough analyses of the interplay of the technical infrastructure and their human users. The presented simulation environment enables assessments of effects such as market frequency, number of participants and social influence on bidding strategies.

In terms of the market server, we are going to implement more innovative mechanisms besides uniform clearing. The energy management gateways are extended in order to better reflect their users' preferences such as comfort, economic

Fig. 21.3 Comparison of clearing price between a market with only one product (15 minutes delivery interval) and a market with two products (5 minutes and 15 minutes delivery time)

performance and social norms. For instance, some users prefer a high degree of comfort, meaning rather small intervals or fixed times when they expect certain applications like dish washers, heat pumps or charging electric vehicles to operate. Others put more weight on a sustainable lifestyle and seek to consume energy when it is available. Nevertheless, such considerations evolve in a social context, and that is why we also investigate effects of social influence on these preferences when individuals interact with each other in their neighbourhood and on social network topologies.

To test the impact of real-world performance, we are currently setting up a distributed environment of 60 Raspberry Pi 3B+ and Zero machines that independently communicate with the market server.

Acknowledgements The research presented in this article was partly funded by the Federal Ministry for Education and Research (BMBF) under contract no "03SFK4F1".

References

1. International Energy Agency (IEA): Medium-Term Renewable Energy Market Report 2015. Market Analysis and Forecasts to 2020 (2015)
2. Niesse, A., Lehnhoff, S., Troschel, M., Uslar, M., Wissing, C., Appelrath, H.-J., Sonnenschein, M.: Market-based self-organized provision of active power and ancillary services. An agent-based approach for smart distribution grids. In: Complexity in Engineering (COMPENG), 2012. 11–13 June 2012, Aachen, Germany; Second Edition of the IEEE Workshop on Complexity in Engineering, pp. 1–5. IEEE, Piscataway (2012)
3. Ringler, P., Keles, D., Fichtner, W.: Agent-based modelling and simulation of smart electricity grids and markets – a literature review. Renew. Sust. Energ. Rev. **57**, 205–215 (2016)
4. Rogers, A., Jennings, N.: Intelligent Agents for the Smart Grid. PerAda Magazine (2010)

5. Sensfuß, F., Genoese, M., Ragwitz, M., Möst, D.: Agent-based simulation of electricity markets -a literature review. Energy Studies Rev. **15**, (2007)
6. Weidlich, A., Veit, D.: A critical survey of agent-based wholesale electricity market models. Energy Econ. **30**, 1728–1759 (2008)
7. Krebs, F.: An empirically grounded model of green electricity adoption in Germany: calibration, validation and insights into patterns of diffusion. J. Artif. Soc. Soc. Simul. **20**(10), (2017)
8. Sorda, G., Sunak, Y., Madlener, R.: An agent-based spatial simulation to evaluate the promotion of electricity from agricultural biogas plants in Germany. Ecol. Econ. **89**, 43–60 (2013)
9. Zhang, T., Nuttall, W.J.: Evaluating government's policies on promoting smart metering diffusion in retail electricity markets via agent-based simulation. J. Prod. Innov. Manag. **28**, 169–186 (2011)
10. Dallinger, D., Wietschel, M.: Grid integration of intermittent renewable energy sources using price-responsive plug-in electric vehicles. Renew. Sust. Energ. Rev. **16**, 3370–3382 (2012)
11. Kahrobaee, S., Rajabzadeh, R.A., Soh, L.-K., Asgarpoor, S.: Multiagent study of smart grid customers with neighborhood electricity trading. Electr. Power Syst. Res. **111**, 123–132 (2014)

Chapter 22
Looking into the Educational Mirror: Why Computation Is Hardly Being Taught in the Social Sciences, and What to Do About It

Wander Jager, Katarzyna Abramczuk, Agata Komendant-Brodowska, Anna Baczko-Dombi, Benedikt Fecher, Natalia Sokolovska, and Tom Spits

Abstract In the current digital era, with an increasingly complex and turbulent society, demand is rising for social scientists capable of analysing behavioural dynamics. Studying behavioural dynamics is a valuable lens, both in public policy making and community planning and in scientific projects on how human behaviour affects ecosystems. Computational social science (CSS) offers a framework for this type of studies, as it connects a complex networked systems perspective with a suite of computational tools and methodologies. Despite its potential and fast growth, CSS is still hardly found in programmes at bachelor and master levels in Europe. We would like to take a closer look at this discrepancy. We discuss why there is a need to develop computational education in the social sciences and why it is a challenge. We consider two perspectives. In the students' perspective, there is the problem of mathematical anxiety and gender inequalities in STEM education. In the academic teachers' perspective, there is the problem of a split within social sciences, inadequacy of statistical models for analysing dynamics, and insufficient educational resources at an appropriate level of difficulty. We build a case for creating a MOOC programme addressed specifically to social sciences students to fill in the existing gaps, facilitate the organisation of learning communities, and advance computational thinking within social science. Creating such a MOOC programme is the goal of our Erasmus+ project "Action for Computational Thinking in Social Sciences" (ACTISS).

W. Jager (✉) · T. Spits
University of Groningen, University College, Groningen Center for Social Complexity Studies, Groningen, The Netherlands
e-mail: w.jager@rug.nl

K. Abramczuk · A. Komendant-Brodowska · A. Baczko-Dombi
University of Warsaw, Institute of Sociology, University of Warsaw, Warsaw, Poland

B. Fecher · N. Sokolovska
The Alexander von Humboldt Institute of Internet and Society (HIIG), Berlin, Germany

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_22

Keywords Computational social science · Computational thinking · Higher education · Educational innovation · Massive open online course (MOOC)

22.1 The Need for Computational Education in the Social Sciences

In the current digital era, with an increasingly complex and turbulent society, demand is rising for social scientists capable of analysing behavioural dynamics. Studying behavioural dynamics is a valuable lens, both in public policy making and community planning, as in scientific projects on how human (economical) behaviour affects ecosystems. According to EU Commission documents [1], "the knowledge economy needs people with the right mix of skills: transversal competences, e-skills for the digital era, creativity and flexibility and a solid understanding of their chosen field (such as in Science, Technology, Engineering and Maths)".

Computational social science (CSS) partly answers this need for skills by offering a framework that connects a complex networked systems perspective with a suite of computational tools and methodologies. CSS covers collecting, processing, and visualising large data sets and simulating the dynamics of behaviour, ranging from processes within the individual to society at large. It provides an interdisciplinary approach to the social sciences [2] and allows for a more precise testing of behavioural mechanisms in a dynamic setting [3]. Thus, it offers tools to address the current problems in the social sciences regarding lack of replication, artificial and non-representative laboratory settings, and the fundamental limitations of statistical analysis [4]. Furthermore it allows linking to other formal models, such as ecological and economical models. This contributes to bringing social scientific knowledge into projects addressing multidisciplinary challenges, such as adaptation to climate change, opinion dynamics, transitions in energy use and food consumption, and the adaptive capacity of societies dealing with migration (see, e.g. [5]). Computer simulations of artificial societies are increasingly being used as a tool to systematically explore the complex behavioural dynamics in social systems [6, 7].

Whereas all of the above leads to an increasing demand for computational social scientists, both public and private employers, including in research intensive sectors, increasingly report mismatches and difficulties in finding the right people for their evolving needs. For example, PhD students and postdocs working on SCC projects usually origin from outside the social sciences. Despite its potential and fast growth, CSS is still hardly found in programmes at bachelor's and master's levels in Europe, be it social psychology, cognitive psychology, or sociology. Hence, most students are not aware of developments in modelling societal dynamics, automated information extraction (big data), and modelling the complexity of individual behaviour and cognition. It seems that – in contrast to research – teaching social sciences students' computational thinking is lagging a bit behind. In other words, in the current educational landscape, there is a missing (or insufficient) piece, and there is still a lot to do in order to meet the increasing demand for computational social scientists. Integrating CSS tools into the curricula would help to enhance problem-solving capacities of social scientists but also offer the necessary approaches in dealing with new kinds of data that are available and relevant for analysing behavioural dynamics, for example, automatically extracted data that allow us to observe individual behaviour, as well as interactions on a large scale of extended periods of time [8].

22.2 Why Are We Less Successful in Education Than We Hoped for?

Looking in the mirror, we may not like to see that CSS teaching has not developed the muscle that we had hoped for. In a survey among 4026 teachers of methods and/or statistics in social sciences (31 per cent of which cover big data analytics or data science methods in their courses), the level of programming and statistical knowledge that the students possess was indicated as the biggest problem [9]. A key question we face is why is CSS – despite its relevance – not being widely taught in social scientific educational programmes. Two main reasons we identify are (1) students feeling uncomfortable with more computational approaches and (2) teaching staff having only limited experience and resources to teach CSS and limited support from managing boards.

Concerning the student perspective, many students of social sciences do not feel capable enough to study subjects which require competences in mathematics and/or ICT. In a typical social studies curriculum, students attend obligatory courses in statistics and quantitative methods. These are usually perceived as difficult, and computational social science is often associated with bad experiences with those courses. This problem is described in the literature as “mathematics and statistics anxiety” [10–12]. Several studies about social sciences students show a high prevalence of statistics anxiety among them (75–80%). This has an impact on students’ subsequent educational and professional paths (e.g. [12]). Considering that CSS relies on programming skills, it is obvious that this anxiety for computing may constitute a first barrier to address.

This problem is related to gender composition of population of social sciences students. Girls are often conditioned from an early childhood to doubt their STEM[1]-related abilities. As a consequence, perceived self-efficacy in advanced ICT and mathematical skills for girls is significantly lower than for boys [13]. At the same time in all the European countries, a student of social sciences is typically a woman. Differences in gender shares of men and women in this group reach even 23 pp. [14]. It is a manifestation of disciplinary segregation, which creates an impression that

[1] Science, technology, engineering, and math

STEM and SSHA[2] are mutually exclusive and contributes to underrepresentation of women in STEM disciplines [15]. At later stages of education, this restricts educational options for young women.

A second barrier resides in the fact that within social sciences faculties, only few – often junior – scientists are familiar with CSS, and if they are familiar, this does not always translate into CSS education. One reason is pretty simple, and it is the division into two almost separate worlds of researchers: using qualitative and quantitative methods, which does not occur in case of STEM disciplines. Therefore, some researchers do not use quantitative methods and are not familiar with CSS methods. More importantly, traditional social science is relying very much on statistics, whereas CSS is focussing more on dynamics of behaviour. Because it is very difficult to translate, for example, regression equations into agent rules, discussion on a theoretical level is often avoided. We do, however, see an increase in field projects where statistical methods are used in parameterising simulation models. Slowly we observe an increase of publications in reputable social scientific journals. This contributes to the acceptance and uptake of CSS in future research projects. Moreover, together with the simulation models that are often available, this provides educational material for societal-relevant SSC projects and classes.

Summarising, to stimulate CSS teaching at social scientific faculties, we need to address computational anxiety in students and support teaching staff with an easily available high-quality material tailored to the needs of students of social sciences.

22.3 How to Change That

A growth can be observed in the number of projects that use CSS as a methodology. These projects are often interdisciplinary, bringing in computational methodologies in domains of, e.g. ecological systems and planning. Hence, an increase is expected in the numbers of PhD students and postdocs that are experienced with these methodologies. However, they are often located outside of social scientific faculties. Using methodologies that are still quite uncommon for many regular social scientists, these young scientists often face a hard time when they aspire a position within a social scientific faculty. It is also worth noting that CSS are also facing a challenge related to its usually interdisciplinary character. Research that is not strictly connected to one discipline but crosses the boundaries (e.g. between sociology, psychology, and educational sciences) can be more difficult to publish, and it also translates to worse access to funding. It has been argued that interdisciplinary character is often a disadvantage when it comes to research evaluation and this claim is supported by qualitative and quantitative studies on research policy (e.g. [16, 17]). If so, researchers may be discouraged to choose a

[2]Social sciences, humanities, and arts

CSS-centred path in their academic career which also leads to a lowered motivation to engage in teaching CSS.

On the other hand, when students become aware of the capacities and demand for CSS methods, they may not be only curious, but may also consider this as an opportunity to develop skills that are increasingly in demand. To help the social scientific community in absorbing the CSS approach, it would be good to offer educational material that supports students in getting an accessible introduction into CSS. This material could be studied independently, but preferably local learning communities emerge around this educational material, especially when exercises address programming issues. Especially here peer support is important, and universities could support such learning communities by allocating support staff.

Much material is already freely available on the web, facilitating the organisation of learning communities. For example, within NetLogo [18], a rich library of exemplary models is available, organised along different scientific disciplines, including the social sciences. NetLogo offers a number of standard models that are inviting to play with, which naturally raise an interest in the computations in the model and its dynamic outcomes. The CoMSES computational library[3] offers a growing collection of executable models that have been used in published research. Many videos can be found on the web explaining computational methods and platforms. To connect with this material, massive open online course (MOOC) modules can be developed addressing different theoretical concepts, applied cases, methodology, and tools. We aim at creating such a modular MOOC within the project "Action for Computational Thinking in Social Sciences" (ACTISS), co-funded by the Erasmus+ Programme of the European Union.

For this endeavour to be successful, several conditions should be met. First, it is especially important that educational materials are designed specifically for students of social sciences. For example, they should not require prior knowledge of advanced maths and should rely on examples that are relevant for social problems. Second, in a modular MOOC, we can bring together videos with basic explanation of a phenomena, papers with the full details, models to exercise with, and a worked-out learning path to build up and test students' knowledge and experience. This opens up the possibility for students to follow more individual learning lines. Third, we should give students the opportunity to get a better understanding of future applicability of computational models in different professional fields. By presenting them real-world cases from public policy, marketing, economy, finances, or others, where computational models are being used to analyse or improve problem-solving capacities in general, a MOOC can inspire students to deepen their knowledge in CSS. Offering a basic introduction to CSS in a comprehensive MOOC for students could be the first step of a broader strategy leading to an increased use of computational thinking in social sciences. By gaining the first experience in CSS and realising the benefits for own research and studies, as well as by identifying

[3]CoMSES computational library: https://www.comses.net/codebases/

the potential to provide enhanced problem-solving capacities and the connectivity to real-world problems, social science students might develop a greater interest in acquiring CSS skills in the future.

At the same time, for teaching staff, it becomes easier to organise various courses when such educational materials are available. MOOCs in particular fit well in new educational formats such as community learning and flipping the classroom. Working in small groups on exercises, watching and discussing the videos on social phenomena, and studying and reflecting on papers contribute to a vivid learning environment where students help each other in mastering the computational skills.

MOOC is an educational format that is well defined and that can reach large audiences – e.g. on FutureLearn platform, dedicated for such courses, an average MOOC has approx. 5000 participants in 1 edition, and it is possible to have multiple reruns of such courses. When we look in the mirror, we may see more possibilities to make our work accessible for various types of learners. Perhaps we could develop exciting computational challenges for high school projects or even primary education. And hopefully they may witness CSS reaching a scientific adulthood.

References

1. EU Commission document: Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the regions supporting growth and jobs – an agenda for the modernisation of Europe's higher education systems. http://eur-lex.europa.eu/legal-content/GA/ALL/?uri=CELEX:52011DC0567. Accessed 21 March 2018 (2011)
2. Cioffi-Revilla, C.: Introduction to Computational Social Science. Principles and Applications. Springer, London (2014)
3. Vallacher, R.R., Read, S.J., Nowak, A. (eds.): Computational Social Psychology. Routledge, New York (2017)
4. Open Science Collaboration: Estimating the reproducibility of psychological science. Science. **349**(6251), (2015). https://doi.org/10.1126/science.aac4716
5. Jager, W.: Enhancing the realism of simulation (EROS): on implementing and developing psychological theory in social simulation. J. Artif. Soc. Soc. Simul. **20**(3), 14 (2017). https://doi.org/10.18564/jasss.3522
6. Gilbert, N., Troitzsch, K.: Simulation for the Social Scientist, 2nd edn. Open University Press, Maindenhead (2005)
7. Squazzoni, F., Jager, W., Edmonds, B.: Social simulation in the social sciences: a brief overview. Soc. Sci. Comput. Rev. **32**(3), 279–294 (2014)
8. Watts, D.J.: Should social science be more solution-oriented? Nat. Hum. Behav. **1**, 0015 (2017)
9. Metzler, K., Kim, D.A., Allum, N., Denman, A.: Who is Doing Computational Social Science? Trends in Big Data Research (White Paper). SAGE Publishing, London (2016). https://doi.org/10.4135/wp160926
10. Ashcraft, M.H.: Math anxiety: personal, educational, and cognitive consequences. Curr. Dir. Psychol. Sci. **11**(5), 181–185 (2002)
11. Onwuegbuzie, A.J., Wilson, V.A.: Statistics anxiety: nature, etiology, antecedents, effects, and treatments–a comprehensive review of the literature. Teach. High. Educ. **8**(2), 195–209 (2003)
12. Zeidner, M.: Statistics and mathematics anxiety in social science students: some interesting parallels. Br. J. Educ. Psychol. **61**(3), 319–328 (1991)

13. Fraillon, J., Ainley, J., Schulz, W., Friedman, T., Gebhardt, E.: Preparing for life in a digital age: The IEA International Computer and Information Literacy Study international report. https://link.springer.com/content/pdf/10.1007/978-3-319-14222-7.pdf. Accessed 21 March 2018 (2014)
14. Eurostat database: European Commission. https://ec.europa.eu/eurostat/data/database. Accessed 21 March 2018 (2017)
15. Buschor, C.B., Berweger, S., Keck Frei, A., Kappler, C.: Majoring in STEM—what accounts for women's career decision making? A mixed methods study. J. Educ. Res. **107**(3), 167–176 (2014)
16. Nightingale, P., Scott, A.: Peer review and the relevance gap: ten suggestions for policy-makers. Sci. Public Policy. **34**(8), 543–553 (2007)
17. Rafols, I., Leydesdorff, L., O'Hare, A., Nightingale, P., Stirling, A.: How journal rankings can suppress interdisciplinary research: a comparison between innovation studies and business & management. Res. Policy. **41**(7), 1262–1282 (2012)
18. Wilensky, U.: Center for Connected Learning and Computer-Based Modeling. Northwestern University, Netlogo (1999)

Chapter 23
Effects of Heterogeneous Strategy Composition on Cooperation in the Repeated Public Goods Game

Marie Lisa Kapeller, Georg Jäger, and Manfred Füllsack

Abstract We investigate a heterogeneous environment of strategic choice in the modelled repeated public goods game. Our objective is to assess the emergence and resilience of cooperation in respect to the coexistence of competing strategies while engaging in cooperation influencing interactions and possible change of strategic behaviour. The chosen strategies follow a simple design, mimicking human behaviour choices while being of strongly reduced complexity. Engaging in different scenarios, we show strategy type distributions where cooperation is stable while for other distributions, specific decay patterns emerge that lead to an overtake of defective strategies. Moreover, we use local effect measurements to examine an individual strategy impact on its environment and how positively or negatively a strategy type behaves. In the light of social contagion, this holds information on how positive or negative a strategy type is for the collective cooperative behaviour.

Keywords Cooperation · Public good · Networks · Agent-based model · Heterogeneity

23.1 Introduction

Heterogeneity is a key factor for many successful populations [1, 5, 11, 21, 23]. The combination of population shares of different strategies can have a positive effect on the achievement of population goals, such as hunting and sharing in packs of predators [17], cooperative nesting and breeding in birds [20], the production of open-source software and refraining from downloading music from the internet [13, 18]. Generally, the coexistence of cooperators and defectors is a typical outcome of most social dilemmas, resulting in a mixed equilibrium state [2].

M. L. Kapeller (✉) · G. Jäger · M. Füllsack
University of Graz, Graz, Austria
e-mail: marie.kapeller@uni-graz.at

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_23

A particular important issue in modern times is the handling of public goods or common goods [7]. Public goods, which are non-excludable and non-rivalrous, can be found in situations ranging from family issues to global warming [9, 12, 19]. However, evolutionary game theory predicts that without additional mechanisms such as punishment or reward [8, 22], the temptation to forgo the public good mostly wins over collective cooperation.

In this study, we investigate the effect of strategy composition on the resilience and sustainability of cooperation in a networked public goods game (PGG). Due to interaction and switching to more successful strategies, networked public good games are highly dynamic, and the initial strategy mix and average cooperation are bad indicators for the overall success of the group in the long run. Thus we use an interdisciplinary approach of complexity research, game theory, agent-based modelling and social sciences in order to simulate the time development of the PGG, both in terms of average cooperation and in terms of strategy composition. We study different initial strategy distributions and find out which mixes are sustainable and in which the cooperation collapses over time. Additionally we introduce measurements to gather information about how each strategy influences its neighbours, giving more insight into the reason behind a collapse or a stable system.

There are several studies that investigate the evolution of the willingness to cooperate in repeated cooperation games. In [10] the prisoner's dilemma was expanded from a two-person to an n-person game. The goal of this investigation was to find the most successful strategies in a multiplayer environment using a tournament setup. In contrast to that, our focus is not the success of individual players, but the sustainability of cooperation in different strategy mixes. A recent study [23] investigated the evolution of cooperation in an evolving PGG. However, the strategies investigated there are simple instantaneous reactions to the behaviour of other players, while the memory of our model makes it possible to investigate more sophisticated strategies. Additionally, we are not limited to groups of players playing against other groups, but our network approach enables us to simulate more complex connections between agents, allowing for a deeper understanding of the evolution of cooperation in various systems.

This paper is organized as follows: In Sect. 23.2 we introduce the model used to simulate cooperative behaviour. Our results are presented in Sect. 23.3 and concluded in Sect. 23.4.

23.2 Model

We use a software modelling approach to create a population of decision-makers playing the repeated public goods game. The model consists of N software agents, who have to decide how much they want to contribute into a common pot. Players are positioned on a network indicating their social contacts (nearest neighbours or link neighbours), which have an influence on individual decisions and ultimately shaping the spread of cooperative and defective behaviour. We use a toroidal lattice topology, where all players are connected with four link neighbours.

Each round, the sum of investments is multiplied by an enhancement factor; afterwards, the resulting dividend is equally divided among all players, irrespective of their invested amount. The payoff of an individual i is given by

$$p_i = G/N - I_i \tag{23.1}$$

with I_i being the individual's investment and the common good

$$G = f * \sum_j I_j \tag{23.2}$$

consisting of the sum of all investments multiplied by an enhancement factor $f > 1$.

In that sense, all the players in the network play the PGG together. However, the topology is important when they make their decision: They can only observe the behaviour of their direct link neighbours, so when trying to judge the overall cooperation or defection status of the whole group, they have only limited information. Each player knows how much their link neighbours invested last turn and the dividend they received from the PGG and update their own investment with regard to their strategy based on this information each round. With $I_i \in [0, 1]$, an investment choice is cooperative if $I_i \geq 0.5$; otherwise the individual chose to defect.

23.2.1 Strategy Types

The development of different strategy types is based on well-known strategies of cooperation games [3, 6, 14–16]. Strategies range from very cooperative to very defective orientations. Individuals of the same strategy type form a population share. Cooperation updates of each strategy type are listed in the following:

- UNCONDITIONAL-COOPERATOR (UC) invests maximum each round,
- UNCONDITIONAL DEFECTOR (UD) does not invest in any round,
- RECIPROCATIVE (R) updates investment according to the mean of their neighbourhood's investment of the previous round (related to tit-for-tat),
- CONDITIONAL COOPERATOR (CC) updates investment according to the PG, increases investment if the received dividend is sufficiently large and decreases investment if dividend is below a threshold,
- ENVIRONMENTALIST (E) updates investment according to the PG, selfishly decreases investment if the PG is sufficiently large and voluntary increases investment when the PG is below a threshold,
- VINDICTIVE (V) starts with full cooperation but loses permanently in its willingness to cooperate for each neighbour that defects even once (related to Friedman strategy).

23.2.2 Evolution

Starting in an optimistic situation, the initial investment for all strategies (except for UD) is very high. Following the game dynamics, agents invest each round according to the investment rule of their strategy. Investment updates are limited to 1% of the population, chosen randomly after each round. If the received payoff of an individual is negative, $p_i < 0$, these individuals have a 0.1% chance to change their current strategy. If players switch their strategy, the new strategy is chosen from strategies of their individual link neighbours who have a different strategy type.

23.2.3 Measurements

Regarding heterogeneous populations, the impact on the collective cooperation state varies for different strategy types. While the global system state is given by the dividend, local activities of each individual are important for social contagion effects, leading to emergent collective behaviour patterns. To quantify the impact of different strategy types, we use two measurements to assess the local influences:

(i) Local mean investment (mean investment of a player's neighbourhood)

$$lmi_j = \frac{1}{4}\sum_{k=1}^{4} I_j^{(k)} \tag{23.3}$$

with I_j^k being the investment of the kth link neighbour of the player j.

(ii) Local investment gap (difference in investments of a player and its neighbourhood)

$$lig_j = I_j - lmi_j, \tag{23.4}$$

with I_j being the investment of the player j.

To analyse typical impacts of each strategy, we denote the average lmi of all individuals of a strategy s as $LMI^{(s)}$ and the average lig of all individuals of a strategy s as $LIG^{(s)}$.

23.2.4 Simulations

We investigate the time evolution of the behaviour of the strategy types by analysis of the (i) mean investment of each population share, (ii) population share sizes, (iii) local mean investment and (iv) local investment gap of each population share. We use a population size of $N = 10{,}000$ individuals, randomly positioned on a toroidal lattice network. The setup leads to an initial situation of high overall investment ($I = 0.98$, std $= 0.02$) of the different population shares, except for unconditional strategies which have an initial investment of $I = 1$ (UC) and $I = 0$ (UD), respectively. The enhancement factor of the public goods game is $f = 1.6$, and the threshold for the investment update of the PG oriented strategies (CC and E) is set at 40% of the maximal possible PG.

Behavioural pattern formation in heterogeneous populations is typically intertwined with population share ratios of the initial configuration. By implementing a small chance for a strategy type change, this impact of initial population share sizes is minimized as well, such that a wide range of possible initial distributions leads to similar outcomes in the collective cooperation. Generally, scenarios ranging from stable collective cooperation to cooperation breakdown are possible. To address differences in the development of collective behaviour, four different scenarios of initial population share sizes are investigated, which are outlined in Sect. 23.3.

A few typical aspects of the cooperation dynamics can be discussed on the simple case of a population consisting mainly of Reciprocative players ($N_R = 5000$) and two equally large population shares of Unconditional-Cooperators ($N^{(UC)} = 2500$) and Unconditional-Defectors ($N^{(UD)} = 2500$). As shown in Fig. 23.1, the collective cooperation gradually declines due to the decrease of UC individuals. Since strategy change from UC to R and UD is the prevalent change, an increase of the population shares of R and UC is visible up to $t \simeq 2500$. The increase of R individuals is followed by a decline until all population shares are equilibrated and total defection has spread through the entire network (Fig. 23.1 (top)). While the collective cooperation declines, the positive impact on a player's neighbourhood declines likewise. Figure 23.1 (bottom left) depicts this phenomenon that with decreasing share of cooperative players, the positive impact of the UC individuals $LIG^{(UC)}$ also decreases, even if they continue to invest the maximal amount into the PG.

Regarding R individuals in a situation where cooperation fails, this population share does not die out, even though some individuals engage in cooperation in the beginning. Due to their investment dynamics, this population share serves as a good transmitter of social contagion processes, as shown in Fig. 23.1 (bottom right), where the transient of the initial positive state is towards no investment gap between R individuals and their neighbourhood. However, without the positive influence of UC, defection spreads through the R population share.

23.3 Results

Our main focus is to examine the impact of strategy type composition on the collective cooperative behaviour. To understand this connection, we tested different scenarios by variation of initial population share sizes. When combining the presented six strategy types (general case) in one population and testing different ratios, we found two main scenarios of cooperation evolution, one leading to

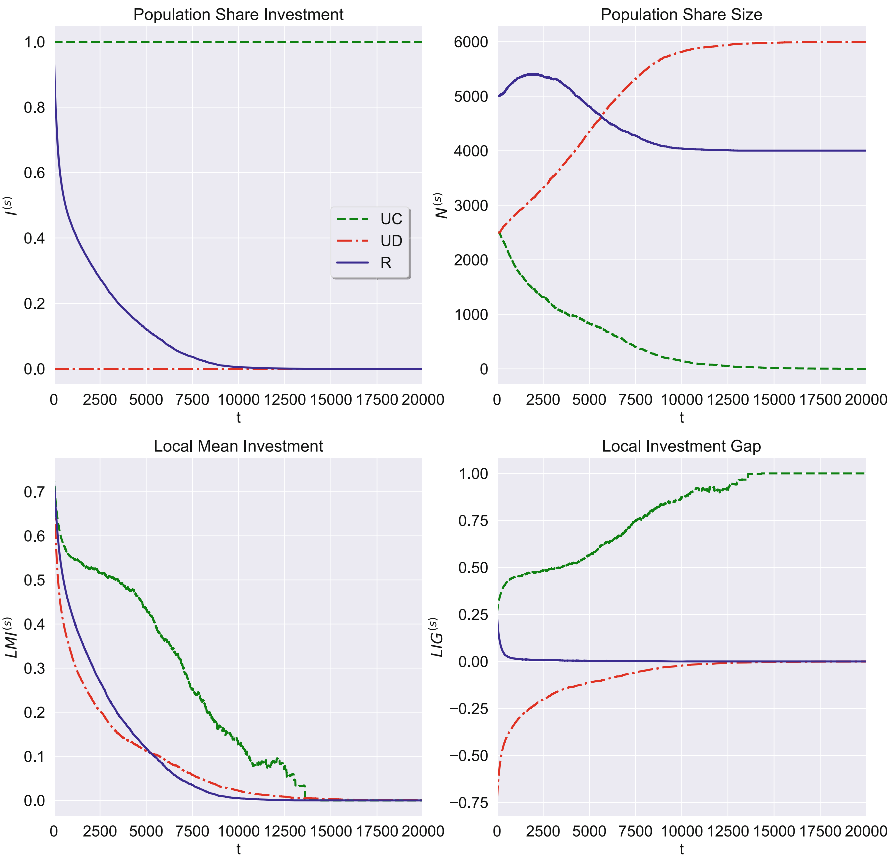


Fig. 23.1 Cooperative behaviour of strategy population shares UC, UD, and R: (tl) mean investment, (tr) population share size, (bl) local mean investment $LMI^{(s)}$, and (br) local investment gap $LIG^{(s)}$ for all three population shares s. Initial distribution of $N^{(R)} = 5000$, $N^{(UC)} = 2500$, $N^{(UD)} = 2500$, population $N = 10{,}000$

cooperation failure and the other to stable cooperation. Moreover, we show a scenario with less strategy types to bring light to cooperative behaviour under direct competition between opposing strategies.

23.3.1 General Case

Regarding general scenarios of mixing the six population shares UC, UD, R, CC, E, V, an essential observation of the decay patterns is the intersection of the decay by a short plateau, where collective cooperation is nearly stable. The position of this plateau strongly correlates to the initial population share sizes. If the formation

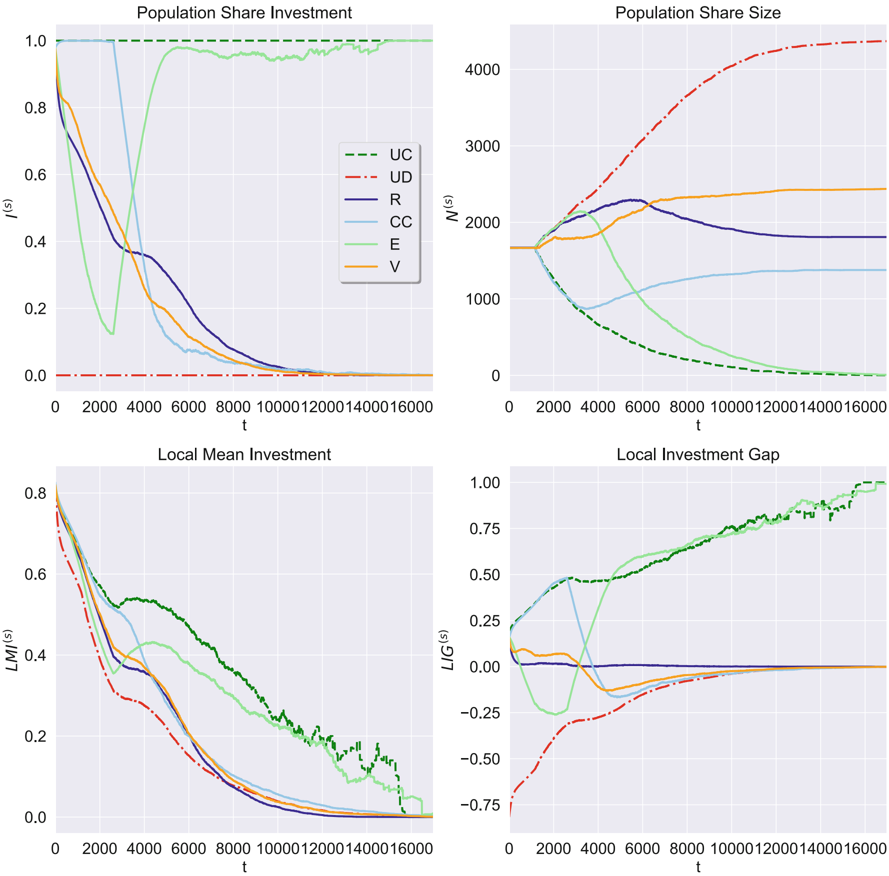


Fig. 23.2 Unstable cooperation: effects of typical decline pattern of cooperative behaviour shown for (tl) mean investment, (tr) population size, (bl) local mean investment, (br) local investment difference for each population share UC, UD, R, CC, E, and V. Initial conditions: equal population share sizes $N^{(s)} \simeq 1666$, population $N = 10{,}000$

of the plateau is near or below 50% of the maximal investment, after a short duration, the cooperation breaks down, and subsequently the collective cooperation is approaching a non-cooperation state. If the initial cooperation decay is less strong, instead of the formation of a cooperation plateau, the cooperation is stable.

Figure 23.2 shows the cooperation failure for equally large initial population share sizes $N^{(s)} \simeq N/6$. Decay patterns of the collective population state follow the decay pattern of the Reciprocative population share closely (Fig. 23.2 (top left, blue line)), while other strategies show a unique development. While the initial decrease in investment of the Environmentalist population share is even more rapid than the investment decrease of the whole population, the following increase causes the formation of the cooperation plateau and, thus, delays the cooperation collapse. In correlation to their cooperative behaviour, after a first increase in the population

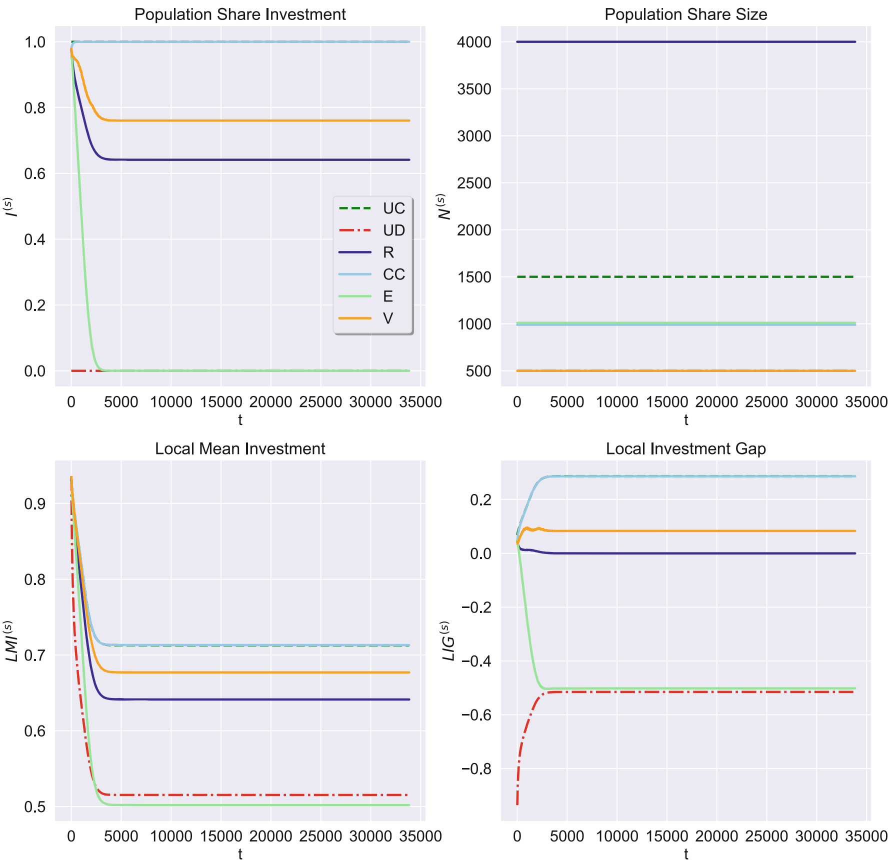


Fig. 23.3 Stable cooperation: example of non-vanishing cooperation shown for (tl) mean investment, (tr) population size, (bl) local mean investment, (br) local investment difference for each population share UC, UD, R, CC, E, and V. Initial conditions: $N^{(R)} = 4000, N^{(UC)} = 1500, N^{(UD)} = 500, N^{(CC)} = 1000, N^{(E)} = 1000, N^{(V)} = 500$, population: $N = 10{,}000$

share of strategy E, this population is, next to Unconditional Cooperators, the only strategy that dies out when cooperation fails.

Figure 23.3 shows a different initial composition of strategy types where cooperation becomes stable shortly after the first decay. Here, strategies are constant since all individuals receive a sufficiently high dividend and obtain a positive payoff.

23.3.2 *Special Case*

To outline direct competition of two opposing strategy types, we investigate populations consisting of Unconditional Cooperators, Environmentalists, and Reciproca-

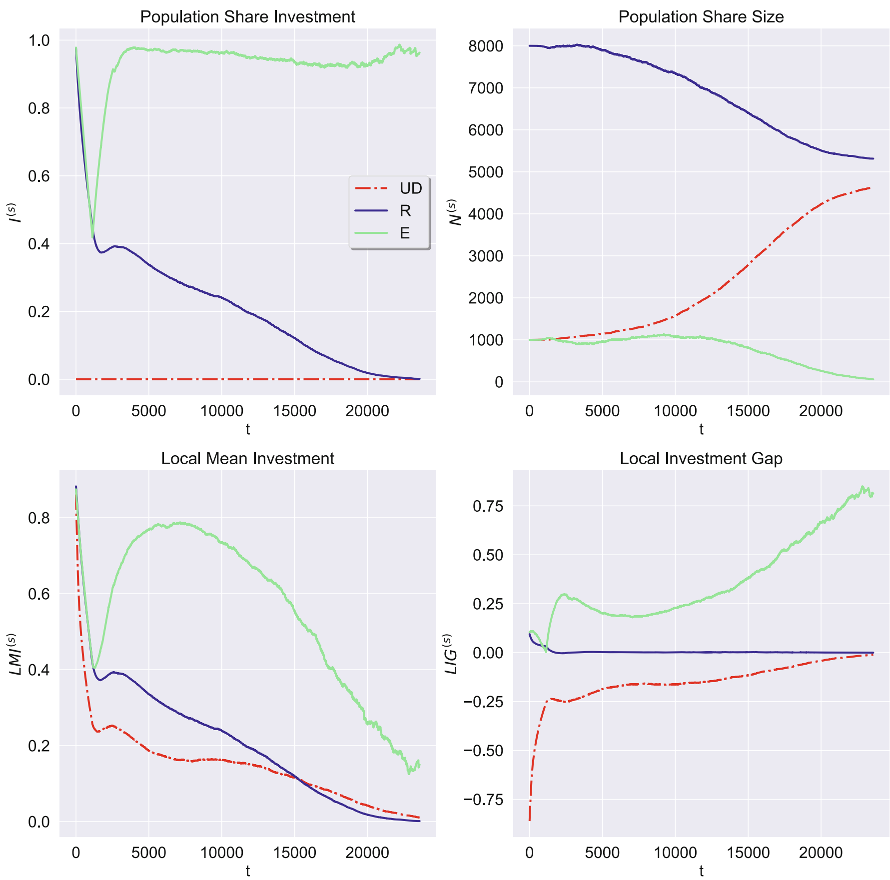


Fig. 23.4 Special case: example of Environmentalists competing against Unconditional-Defectors shown for (tl) mean investment, (tr) population size, (bl) local mean investment, (br) local investment difference for each population share. Initial population share sizes of $N^{(R)} = 8000$, $N^{(UD)} = 1000$, $N^{(E)} = 1000$ and population size of $N = 10{,}000$

tives, the latter serving as transmitter of social contagion processes. Figure 23.4 shows the results for an initial distribution of 10% UD and E each. Here, the cooperation plateau is transformed into a short cooperation increase in the R population, which again closely reflects the mean cooperation of the population. Similar to the general case, cooperation is lost after this short delay and the population of Environmentalists goes extinct. Regarding the local measurements, the fast increase of cooperative behaviour of Environmentalists is reflected by the $LMI^{(E)}$ but is not visible in the $LIG^{(E)}$, which peaks shortly before this strategy type dies out.

23.4 Conclusion

In this study we presented a model to investigate the time development of cooperation in a public goods game. Using heterogeneous populations, we observe, next to the collective cooperation, the non-trivial development of the shares of each strategy. Moreover, we provide local effect measurements that give insight into positive or negative influences of each strategy type on their direct neighbourhood.

An essential result of our investigation on strategy composition was the partition in successful populations, which show stable cooperation above 50% overall cooperation, and unsuccessful populations with complete cooperation loss. We found that the crucial point in the time development is the cooperation plateau that forms after a while. If cooperators still benefit during this plateau, cooperation will most likely remain stable. However, if cooperators experience a net loss, they will start looking for alternative strategies and find that less cooperative strategies offer an advantage. This information is transmitted through the network, and after a while none of the players invest anything anymore and a Nash equilibrium is reached.

Another important aspect we observed was the different manifestation of the cooperative behaviour and the underlying strategy. While cooperative behaviour of individuals might be similar in a nonthreatening situation, this behaviour can diverge when tension in the collective arise. Thus, gaining knowledge on underlying strategies of behaviour to cope with dissimilar situations might be fundamentally important for systems approaching unstable regions. We believe that the willingness to cooperate is only a small part of the complete information, since the time development of certain sustainability measurements can be completely different from the time development of the willingness to cooperate itself.

When investigating processes involving social contagion, an important aspect is the local influence of a given behavioural pattern. We captured these effects by using the local mean investment and local investment gap as measurements to distinguish strategy type impacts. Here, we find that the time development of a potential positive or negative impact can evolve fundamentally different than the investing behaviour. Even though each strategy type produces unique local effects, this correlation is non-trivial such that the current collective cooperative state and the population share size of a strategy are influencing the effective local impact.

Even though the PGG is an abstract game, it shares many features with cooperation behaviour in real life. A prominent example is climate change mitigation, where mitigation efforts have to rely on the willingness of individuals to contribute voluntarily [4]. In that sense, many of the insights we gained could also be translated to development of cooperation in real-life systems.

References

1. Amaral, M.A., Wardil, L., Perc, M., da Silva, J.K.: Evolutionary mixed games in structured populations: cooperation and the benefits of heterogeneity. Phys. Rev. E **93**(4), 042304 (2016)

2. Archetti, M., Scheuring, I.: Coexistence of cooperation and defection in public goods games. Evolution **65**(4), 1140–1148 (2011)
3. Axelrod, R.: Effective choice in the prisoner's dilemma. J. Confl. Resolut. **24**(1), 3–25 (1980)
4. Carattini, S., Levin, S., Tavoni, A.: Cooperation in the climate commons. Rev. Environ. Econ. Policy. **13**(2), 227–247 (2019)
5. Fischbacher, U., Gachter, S.: Social preferences, beliefs, and the dynamics of free riding in public goods experiments. Am. Econ. Rev. **100**(1), 541–56 (2010)
6. Fischbacher, U., Gächter, S., Fehr, E.: Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. **71**(3), 397–404 (2001)
7. Hardin, G.: The tragedy of the commons. Science **162**(3859), 1243–1248 (1968)
8. Helbing, D., Szolnoki, A., Perc, M., Szabó, G.: Punish, but not too hard: how costly punishment spreads in the spatial public goods game. New J. Phys. **12**(8), 083005 (2010)
9. Kollock, P.: Social dilemmas: the anatomy of cooperation. Ann. Rev. Sociol. **24**(1), 183–214 (1998)
10. Manhart, K., Diekmann, A.: Cooperation in 2-and n-person prisoner's dilemma games: a simulation study. Anal. Kritik **11**(2), 134–153 (1989)
11. Meloni, S., Xia, C.Y., Moreno, Y.: Heterogeneous resource allocation can change social hierarchy in public goods games. R. Soc. Open Sci. **4**(3), 170092 (2017)
12. Milinski, M., Semmann, D., Krambeck, H.J., Marotzke, J.: Stabilizing the earth's climate is not a losing game: supporting evidence from public goods experiments. Proc. Natl. Acad. Sci. **103**(11), 3994–3998 (2006)
13. Myatt, D.P., Wallace, C.: Evolution, teamwork and collective action: production targets in the private provision of public goods. Econ. J. **119**(534), 61–90 (2009)
14. Nowak, M.A.: Five rules for the evolution of cooperation. Science **314**(5805), 1560–1563 (2006)
15. Nowak, M.A., Sigmund, K.: Tit for tat in heterogeneous populations. Nature **355**(6357), 250 (1992)
16. Nowak, M., Sigmund, K.: A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. Nature **364**(6432), 56 (1993)
17. Packer, C., Scheel, D., Pusey, A.E.: Why lions form groups: food is not enough. Am. Nat. **136**(1), 1–19 (1990)
18. Palfrey, T.R., Rosenthal, H.: Participation and the provision of discrete public goods: a strategic analysis. J. Public Econ. **24**(2), 171–193 (1984)
19. Peters, H.E., Ünür, A.S., Clark, J., Schulze, W.D.: Free-riding and the provision of public goods in the family: a laboratory experiment. Int. Econ. Rev. **45**(1), 283–299 (2004)
20. Rabenold, K.N.: Cooperative enhancement of reproductive success in tropical wren societies. Ecology **65**(3), 871–885 (1984)
21. Santos, F.C., Pinheiro, F.L., Lenaerts, T., Pacheco, J.M.: The role of diversity in the evolution of cooperation. J. Theor. Biol. **299**, 88–96 (2012)
22. Szolnoki, A., Perc, M.: Reward and cooperation in the spatial public goods game. EPL (Europhys. Lett.) **92**(3), 38003 (2010)
23. Zhang, H., Perc, M.: Evolution of conditional cooperation under multilevel selection. Sci. Rep. **6**, 23006 (2016)

Chapter 24
A Health Policy Simulation and Gaming Model of Ebola Haemorrhagic Fever and Zika Fever

Setsuya Kurahashi

Abstract This paper proposes a gaming and simulation model of Ebola haemorrhagic fever and Zika fever. The mathematical modelling of infectious diseases such as SIR (susceptible, infected, recovered) model and SEIR (S, Exposed, I, R) has been widely used to understand the epidemic of influenza, smallpox, to name a few. Recently, an agent-based model has been adopted to represent the behaviour of each person on the computer. The study designs a model in which health policies are considered such as vaccine stocks, antiviral medicine stocks, the number of medical staff to support overseas and so forth. The infectious disease model of Ebola fever and Zika fever is also implemented in the model. Besides, based on these simulation models, we developed a medical policy decision game that dealt with infections as serious games and verified the effect. As results of experiments using the model, we found that preventive vaccine, antiviral medicine stocks and the number of medical staff are crucial factors to prevent the spread. Also, a modern city is vulnerable to Zika fever due to commuting by train. And it has also been found that self-control and restraint on immigration are essential, in addition to the vaccine reserve amount and the timing of decision-making for medical support to the partner country where the occurrence of infection has spread.

Keywords Infectious disease · Ebola haemorrhagic fever · Zika fever

24.1 Introduction

SIR model and SEIR model have been used to analyse such diseases. These models have a powerful ability to make a real-time estimation. The SIR model, however, has difficulty to decide which measures are effective. The model has a few parameters as an infection rate to represent infectiveness. Therefore, the SIR model has limited

---

S. Kurahashi (✉)
Graduate School of Business Sciences, University of Tsukuba, Bunkyo, Tokyo, Japan
e-mail: kurahashi.setsuya.gf@u.tsukuba.ac.jp

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_24

ability to evaluate the effect of the temporary closing of classes at school. Even though SEIR model has several parameters, it is difficult to make a model of various inhabitants or routes of infection. The agent-based approach or the individual-based approach has been adopted to conquer these problems in recent years [1–11]. The model enables to represent the behaviour of each person. It also reveals the spread of infection by simulation of the contact process among people in the model.

In this study, we developed a model to simulate smallpox and Ebola haemorrhagic fever and Zika fever based on the infectious disease studies using agent-based modelling [12, 13]. This paper builds upon previous work as reported in [14, 15]. What we want to know is how to prevent an epidemic of infectious diseases not only using mechanisms of the epidemic but also decision-making of health policy. Most importantly, we should make a decision in our modern society where people are on the move frequently worldwide, so we can minimise the economic and human loss caused by the epidemic.

24.2 Cases of Infectious Disease

24.2.1 Ebola Haemorrhagic Fever

The Ebola epidemic began in Guinea in Dec. 2013. The authorities of south-west African countries have launched a state committee of emergency. They have taken measures to cope with the situation. In the measures, Guinea adopted the prohibition of entry over the boundary.

24.2.2 Zika Fever

Zika fever is an infectious disease caused by Zika virus through the bite of mosquitoes [16, 17]. Most cases have no symptoms, and present are usually mild including fever, red eyes, joint pain and a rash[18]. The Zika fever can cause microcephaly that crucially affects babies by small head circumference.

24.3 The Health Policy Simulation Model of Infectious Disease

A health policy simulation model of infectious disease was developed based on Burke's model. Our model has functions to simulate smallpox and Ebola haemorrhagic fever. In addition to the previous model, however, inhabitants of our model commute their offices by train every day. It consists of interaction with the agent population. Each interaction results in contact, and then the contact results in a transmission of the infection. It is caused from the contacted agent to the active agent stochastically.

24.3.1 The Base Model of Ebola Haemorrhagic Fever

In the event the active agent contracts the disease, she turns a stage from non-infection to latent infection and her own internal clock of disease progression begins. After 7 days, she begins infecting others. However, her disease is not specified in this stage. After 3 days, she begins to have vomiting and diarrhoea, and the disease is specified as Ebola. Unless the infected individual is dosed with antiviral medicine within 3 days of exposure, the medicine is ineffective. This is an imaginary medicine to play the policy game. At the end of day 12, individuals are assumed to hospitalise. After 4 more days, during which they have a cumulative 90% probability of mortality, surviving individuals recover and return to circulation permanently immune to further infection. Dead individuals are placed in the morgue. Immune individuals are returned to their home. Other settings are the same as smallpox. The upper limit vaccination a day is set at 100, and the coverage rate of trace vaccination is set at 0.7.

24.3.2 The Base Model of Zika Fever

About 80% of cases have no symptoms which are called latent infection, but the latent patients can transmit Zika virus to other mosquitos. The incubation period of Zika virus disease is not clear, which is likely to be 3 to 9 days. After the incubation period, the symptoms including fever, skin rashes, conjunctivitis, muscle and joint pain, malaise and headache occur and last for 6 days. Zika virus disease is relatively mild and requires no specific treatment, so any strategies are not selected in the model [19, 20].

In this model, a mosquito is designed as an agent as well as an inhabitant agent. Zika virus is transmitted to people via the bite of an infected mosquito, the *Aedes* genus. This process is the same as dengue. Therefore, we adopted the process of dengue disease such as the life and habits of mosquitoes that transmit dengue. Mosquitos transmit viruses by bite. The infection is spread at any time of day. Humans are also the primary host of the virus. An infection can be acquired via a single bite. About 10 days later, the virus can be spread to other humans. An adult mosquito can live for 30 days with the virus.

Mosquitoes live around each town, office and school in Zika fever model. The areas they live in overlap with human areas. Therefore the Zika virus can be transmitted between mosquitoes and human. Additionally, mosquitoes also live around a rail station in another Zika fever model. Inhabitants in the model commute their offices by train every day. We are concerned about the possibility

Table 24.1 Summary of Zika infection

Model	Habitat	No infection	Infection Freq.	Max infection
Base	All without station	0: 88%	1–20: 12%	21
Railway	All	0: 77%	1–40: 22%	127
Railway	Station	0: 92%	1–35: 8%	34
Railway	School	0: 91%	1–20: 9%	21
Railway	Town	0: 92%	1–15: 8%	14
Railway	Office	0: 97%	1–10: 3%	3
Railway	All without station	0: 82%	1–35: 17%	61

that mosquitos are infected with Zika virus from people in both towns. Therefore, we conducted an experiment on whether the train station can be the new source of infection or not. A mosquito bites human once per 4 days. An infection rate from a mosquito to human is set at 0.5 and from human to a mosquito is set at 1.0. Table 24.1 shows experimental results of Zika infection model. The worst case of experimental outbreak was the railway model where station, towns, offices and schools are mosquito habitats. When a station habitat was eliminated, the number of infection decreased by half of all habitats case.

24.4 Infectious Disease Medical Policy Game

In the experiment so far, we have seen that it is difficult to prevent the spread of infection in areas with many railway users. A contemporary society faces the risk of the infection spreading across borders like Ebola haemorrhagic fever and new influenza. In this case, international cooperation of medical policy team is important. The World Health Organisation (WHO) is promoting international cooperation for measures against infectious diseases and frequently taken up at the G7 Summit and other conferences. Therefore, in order to expand the model so far, we developed a cooperative game for countermeasures against infectious diseases and conducted an experiment to promote cooperation between two countries or regions.

In this game, it is supposed that a new type of infectious disease occurs in which the disease is similar to smallpox and Ebola haemorrhagic fever. Vaccine and antiviral medicine for the disease are already developed and provided to the market in this model. Players as authorities of two countries decide the amount of both medicine stocks according to their restricted budget. They also need to decide the number of medical staff, blockade and restrictions on outings. The players should consider giving support medicine and staff to countries to prevent or control its epidemic for his/her own country while taking account of economic cost and loss. Travel restrictions have a huge economic impact, while it is very effective in stopping an outbreak. Supporting to another country means decreasing its own preparations. Thus, this game has a complicated structure of trade-offs among cost, effect, cooperation and defence.

24.4.1 Medical Policy Model

The medical policy leaders of both countries will make decisions on the following policies. (1) Number of vaccines ordered for stockpiling (2) Number of antiviral drug orders for stockpiling (3) Employment of medical staff who can respond to Ebola haemorrhagic fever Discussions on the collaboration of infectious disease countermeasures with neighbouring countries are underway, and policies that support each other can be implemented. (4) Number of vaccine support to neighbouring countries (5) Number of antiviral drugs supported in neighbouring countries (6) Number of support for medical staff who can cope with infectious diseases to neighbouring countries However, since this assistance will reduce the stock and medical staff of their country, duty of support is exempted if danger is approaching. On the other hand, as an emergency response, it is possible to instruct the residents to go out and to order restrictions on entry and departure with neighbours. (7) Rate of voluntary ban on leaving home (8) Rate of restrictions on immigration (9) Mass vaccination for all residents (10) Trace vaccination for infected people (11) A dose of antiviral drug to a contact with an infected person (12) A dose of antiviral drug to a contact getting fever with an infected person.

Players get a report of infection status including her/his own country and a neighbouring country every 10 days. Although they make a decision to avoid pandemic in their own country as they cooperate together when they get each report, they have to cope with a difficult situation within a predetermined budget. They are admitted to using additional budget by the government, but the smaller they use the budget, the more advantage they get to win the game. For the sake of their budget management, the following parameters are pre-defined in a game; budget, vaccine price, antiviral drug price, personal expenses of medical stuff, economical cost of a voluntary ban on leaving home and economical cost of restrictions on immigration.

As the first experiment of medical policy game, S town was a little delayed to order enough vaccine and supported C town in too much medicine. Therefore, S town failed to stop its own epidemic. C town quickly decided to order plenty of vaccines, antivirus drug corresponding to the epidemic, and then C town succeeded to stop it. The second experiment in which players experienced the first game showed that S town ordered vaccine twice as much as the 1st game when the epidemic was found. C town decided to send stuff to Square town and a partial voluntary ban on leaving home and blockade. As a result, both towns succeeded to stop the epidemic.

24.5 Conclusion

In this study, we developed the infection process of Ebola haemorrhagic fever and Zika fever as a deductive model based on ABM. Symptoms after infection and each parameter were defined based on actual data, and residents and regional models were defined as two simple regional models as possible. Experimental results showed that trace serum administration is most effective in the case of Ebola haemorrhagic fever, and complete suppression of infection is difficult when railway use is assumed for commuting as public transportation. In addition, as international cooperation by medical authorities is indispensable today, we have constructed a medical policy game that encourages bilateral coordination and conducted experiments. As a result, the experience of the game prompted cooperative action among players. It also gave the opportunity for them to learn the prevention effect by forced policies such as a partial voluntary ban on leaving home and border blockade.

References

1. Burke, S., et al.: Individual based computational modeling of smallpox epidemic control strategies. Acad. Emerg. Med. **13**(11), 1142–1149 (2006)
2. Longini, J.I.M., et al.: Containing a large bioterrorist smallpox attack: a computer simulation approach. Int. J. Infect. Dis. **11**(2), 98–108 (2007)
3. Gilbert, N.: Agent-Based Models, vol. 153. Sage, Los Angels (2008)
4. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press, New York (2010)
5. Okabe, N.: Risk and benefit of immunisation: infectious disease prevention with immunization. Iryo Shakai **21**(1), 33–40 (2011)
6. Epstein, J.M., et al.: Toward a containment strategy for smallpox bioterror: an individual-based computational approach, Center on Social and Economic Dynamics Working Paper No. 31, The Brookings Institution (2002)
7. Epstein, J.M.: Generative Social Science: Studies in Agent-Based Computational Modeling. Princeton University Press, New Jersey (2006)
8. Ohkusa, Y.: An evaluation of counter measures for smallpox outbreak using an individual based model and taking into consideration the limitation of human resources of public health workers. Iryo Shakai **16**(3), 295–284 (2007)
9. Hunder, E., Namee, B.M., Kelleher, J.D.: A taxonomy for agent-based models in human infectious disease epidemiology. J. Artif. Soc. Social Simul. **20**(3), 2 (2017)
10. Yoneyama, T., Das, S., Krishnamoorthy, M.: A hybrid model for disease spread and an application to the SARS pandemic. J. Artif. Soc. Social Simul. **15**(1), 5 (2012)
11. Bruch, E., Atwell, J.: Agent-based models in empirical social research. Sociol. Methods Res. **44**(2), 186–221 (2015)
12. Kuhlman, C.J.: Hybrid agent-based modeling of Zika in the United States. In: Proceedings of the 2017 Winter Simulation Conference (2017)
13. Wiratsudakul, A., Suparit, P., Modchang, C.: Dynamics of Zika virus outbreaks: an overview of mathematical modeling approaches. PeerJ **6**, e4526 (2018)
14. Kurahashi, S., Terano, T.: A health policy simulation model of smallpox and Ebola haemorrhagic fever. In: Jezic, G., Howlett, R., Jain, L. (eds.) Agent and Multi-agent Systems: Technologies and Applications. Smart Innovation, Systems and Technologies, vol. 38. Springer, Cham (2015)
15. Kurahashi, S.: A health policy simulation model of Ebola haemorrhagic fever and Zika fever. In: Jezic, G., Chen-Burger, Y.H., Howlett, R., Jain, L. (eds.) Agent and Multi-agent Systems: Technology and Applications. Smart Innovation, Systems and Technologies, vol. 58. Springer, Cham (2016)
16. Chen, L.H., Hamer, D.H.: Zika virus: rapid spread in the Western hemisphere. Ann. Intern. Med. **164**, 613–615 (2016). https://doi.org/10.7326/M16-0150

17. Hennessey, M., Fischer, M., Staples, J.E.: Zika virus spreads to new areas – region of the Americas, May 2015–Jan 2016. Cent. Dis. Control Prev. **65**(3), 55–58 (2016)
18. WHO: Zika virus, World Health Organization, Media centre, Fact sheets, Jan 2016
19. ECDC: Aedes aegypti – Factsheet for Experts, Health Topics: Vectors: Mosquitos. European Centre for Disease Protection and Control. Retrieved June 2018
20. CDC: Zika Virus, Centers for Disease Control and Prevention, U.S. Department of Health & Human Services, Retrieved June 2018

Chapter 25
A Toy Model for Tertiary Educational Choices in Italy

Silvia Leoni

Abstract This work explores the determinants behind university enrolment decisions, by modelling the choice of young Italians to attend university or to drop out and enter the labour market, by making use of an agent-based model. Determinants behind this choice include expectations on income gathered through interaction with network, as well as peers' influence and perceived effort of education.

Keywords Agent-based modelling · Education · Earnings · Italy

25.1 Introduction

Although the low level of tuition fees and absence of other access barriers, Italy is characterized by low educational attainments at university level.

Within the 10-year European strategy Europe 2020, the European Commission elaborated five targets to be reached by 2020, among which the goal concerning education includes reaching a share of 40% of 30–34-year-old completing a third level education. These common goals translate into national targets for member states. Italy, in particular, aims at reaching 26–27% of people aged between 30 and 34, having attained a tertiary education degree. Indeed, according to Eurostat, in 2017 the 26.5% of Italians aged between 30 and 34 has obtained a higher education degree, placing Italy in line with national targets. Nevertheless, Italy performs badly when compared to the other EU countries: a lower share is registered only in Romania (26.3%).[1]

[1]See [6] for an overview of the tertiary education attainment in the EU countries, as well as of national targets.

S. Leoni (✉)
Department of Economics and Social Sciences, Università Politecnica delle Marche, Ancona, Italy
e-mail: s.leoni@univpm.it

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_25

Moreover, [4] who investigated earnings gaps across generations in Italy, find that the generation born between 1975 and 1979 suffered of a remarkable earnings loss from the first entry in the job market, with respect to the previous generations. In particular, high-skilled workers of the 1975–1979 generation, who completed a tertiary educational level, have been affected by a decrease in their earnings in a much larger measure than low-educated workers, without a catch up of this gap in the following years.

Based on these observations, this work wants to analyse the determinants behind university enrolment decisions by means of an agent-based model and explore whether and in what measure these determinants could explain the low educational attainment characterizing Italy; in particular, a special focus is devoted to the income distribution in order to understand whether the earning loss for high-skilled workers observed by [4] might have discouraged young Italians in pursuing a university degree.

The ABM allows for modelling the decision of individuals to attend university or to drop out and enter the labour market, by taking into account heterogeneity of agents, social interactions and exchange of information, as it is better explained in Sect. 25.3.

25.2 State of the Art

This work contributes to the existing literature in an original way by means of the methodology adopted. The application of ABM in the field of economics of education is relatively new, especially in the evaluation of determinants of schooling decisions. A valid example can be found in [1] who develop an agent-based system in order to model school decisions in French-speaking Belgium; his aim is to assess the impact of school choice on school segregation. The closest example to this work can be found instead in [3]. Using empirical data for France, [3] proves that among the determinants of the distribution of educational choices across social groups, social influence among agents of a network cannot be ignored.

25.3 The Model

25.3.1 Agents

In this model, agents' population is constituted by two generational groups: *juniors* and *seniors*. *Juniors* have completed secondary education and have to decide whether to enrol or not at university. *Seniors* are in the labour market and can be skilled or unskilled, depending on whether they attended university in their previous period. *Juniors* are provided with a monetary endowment deriving from their parental income, which can be spent in education or in the transition from school to the job market. Agents also own a certain *ability*, capturing innate talent and personal skills. One may take into account skills differentiation, by identifying different groups of agents in such a way that the distribution of agents' ability has common variance σ^2 but different mean across groups.

25.3.2 Environment

The agent's set is organized in a social environment with N neighbourhoods representing agent's social relations: family, friends and acquaintances. Agents' population is modelled as a social network structured in social circles. As stated in [2], the idea of social circles is able to embed the key characteristics of real social networks, such as low density, high clustering and assortativity of degree of connectivity, which other network types fail to reflect. Circles help define agents' neighbours as each agent's circumference creates a cut-off limiting the size of personal network and includes all the agents within a given radius as neighbours. Figure 25.1 shows a representation of the social circles structure, drawn from NetLogo.

In order to consider family effect, the model allows to consider stronger ties with parents with respect to the other connections, by assigning a weight w to them, that allows to account for up to double a standard unweighed connection.

To take into account the social context, agents are not randomly located in the world space, but they are placed according to a certain level of initial segregation, so to reflect the possibility that individuals with similar educational level will tend to have friends and acquaintances with their same educational level.

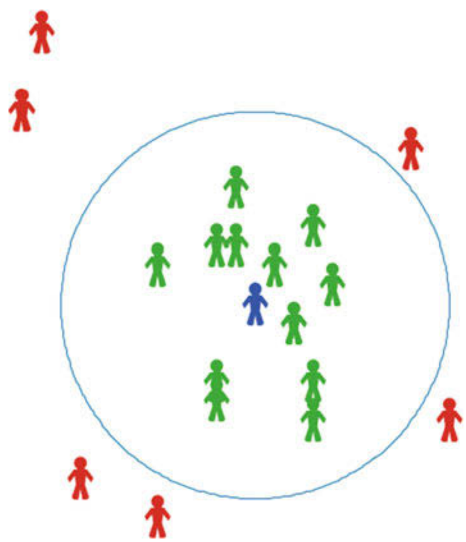


Fig. 25.1 The circle around the blue agent defines her personal network that is represented by the green agents. The red ones outside the reach of the circle radius are not agent's neighbours

25.3.3 Budget Constraint

Juniors will ponder the choice to enrol at university if the following budget constraint is satisfied:

$$X_{i,t} - \text{CostEdu} > 0 \tag{25.1}$$

where $X_{i,t}$ represents the endowment agents are provided with and CostEdu is the cost of education, which is assumed to be exogenous and fixed according to the educational system taken into account.

To replicate real features of the Italian university system, a possible extension of the constraint would be to add the possibility to receive a full scholarship from the government[2] when parental income $Y_{pi,t}$ does not exceed a certain threshold c:

$$Y_{pi,t} < c. \tag{25.2}$$

In this case, the cost of education would be totally funded by the government:

$$\text{Costedu} = 0. \tag{25.3}$$

25.3.4 Consumption

If the budget constraint (25.1) is satisfied, then real expected consumption for skilled workers at time $t + 1$ will be:

$$C^e_{i\text{skilled},t+1} = X_{i,t} - \text{CostEdu} + Y^e_{i\text{skilled},t+1}, \tag{25.4}$$

while real consumption for unskilled workers will be:

$$C^e_{i\text{unskilled},t+1} = X_{i,t} + Y^e_{i\text{unskilled},t+1} \tag{25.5}$$

where $Y^e_{i\text{skilled},t+1}$ and $Y^e_{i\text{unskilled},t+1}$ are, respectively, the average expected income for skilled and unskilled workers.

[2]Scholarships to fund tertiary educational studies are provided by regional bodies devoted to study rights and they can include up to total exemption from tuition fees, daily meals at the university restaurant, free accommodation in a student dorm or a contribution for rent payment. At the entrance at university the scholarship awarding is generally only income-based and requirements are usually related to the level of ISEE, an indicator for the economic situation of households. Reductions or total exemption from tuition fees could also be offered by the university itself depending on the level of ISEE as well as on the final grade obtained with secondary school diploma.

By taking the difference between (25.4) and (25.5), the extra consumption expected to bear when enrolling at university will be:

$$\Delta C^{e}_{i,t+1} = \Delta Y^{e}_{i,t+1} - \text{Costedu}, \tag{25.6}$$

where $\Delta Y^{e}_{i,t+1}$ is the expectation on the extra income made by skilled workers with respect to unskilled workers. These expectations on future income are modelled as naive expectations based on the information set of *senior* neighbours:

$$Y^{e}_{i,t+1} = E_t(Y_{t+1}|\Omega_{n,t}) = \frac{\sum_{i=1}^{nw} wY_{n\text{sen},t}}{nw} \tag{25.7}$$

$$\Delta Y^{e}_{i,t+1} = \frac{\sum_{i=1}^{nw} wY^{S}_{n\text{sen},t}}{nw} - \frac{\sum_{i=1}^{nw} wY^{U}_{n\text{sen},t}}{nw} \tag{25.8}$$

Income is exogenously modelled using data from the Survey on Household Income and Wealth (SHIW) provided by the Bank of Italy every 2 years. Alternatively, the expected extra consumption can be written in the following way and in this case can be interpreted as the percentage of benefits over costs:

$$\Delta C^{e}_{i,t+1} = \frac{\Delta Y^{e}_{i,t+1} - \text{Costedu}}{\text{Costedu}} \tag{25.9}$$

25.3.5 Preference for Enrolling

Preference for enrolling is an additive function, depending on a *materialistic* term represented by expectations on extra consumption, a *social* term SI_{it} and an *intangible* term, given by the disutility for the effort of education:

$$P_{it} = \Delta C^{e}_{i,t+1} + SI_{it} - \text{Effortedu} \tag{25.10}$$

where SI_{it} represents social influence, given by the fraction of the number of peers within one's neighbourhood deciding to enrol over the total number of peers in the neighbour. Its purpose is to reflect a merely imitative behaviour:

$$SI_{it} = \frac{N\text{peersenrolling}}{N_{n\text{peers}}}, \tag{25.11}$$

Effortedu stands for the effort necessary to obtain a university degree. It is assumed to depend only on individual ability through a concave function: with higher individual ability a_{it}, the less the effort to get a university degree.

$$\text{EffortEdu} = (1 - a_{it})^{\gamma} \tag{25.12}$$

where a_{it}, measuring individual ability, as in [5] is assumed to be included between 0 and 1, and $\gamma > 1$ measures the concavity of the function.

The function for the preference for enrolling allows to consider solely the economic motivations guiding the enrolment decision, i.e. the function for the expected extra consumption $\Delta C^e_{i,t+1}$ and/or the intangible motivations identified in the rest of the equation.

Agents enrol at university with a probability increasing in the level of preference P,

$$Pr_{\text{enroll}} = f(P_{it}). \tag{25.13}$$

Building on [3], this could be written as:

$$Pr_{it}(\text{enroll}) = \frac{\exp(P_{it})}{1 + \exp(P_{it})} \tag{25.14}$$

25.4 Potential Results

The model developed in this work explores the factors influencing educational choices at tertiary level and takes into account both economic and social motivations. By considering the ability of agents to interact and adapt their choices to their neighbours, this work could bring light on the decision-making process of young Italians facing the decision to enrol or not at university and help understand whether and in what measure income distribution, family and peers influence, as well as heterogeneity in personal skills could play a role as push or pull factors in the choice considered. The analysis could also highlight the necessity for further elements to be included in the construction of the model, which could better explain low educational attainment levels in Italy.

References

1. Friant, N.: Choix de l'école et ségrégations scolaires dans un espace local en Belgique francophone. Inégalités Éducatives et espaces de vie (2015)
2. Hamill, L., Gilbert, N.: Social circles: a simple structure for agent-based social network models. J. Artif. Soc. Social Simul. **12**(2), 3 (2009)
3. Manzo, G.: Educational choices and social interactions: a formal model and a computational test. In: Class and Stratification Analysis, pp. 47–100. Emerald Group Publishing Limited, Bingley, United Kingdom (2013)
4. Naticchioni, P., Raitano, M., Vittori, C.: La Meglio Gioventù: earnings gaps across generations and skills in Italy. Econ. Polit. **33**(2), 233–264 (2016)
5. Staffolani, S., Valentini, E.: Bequest taxation and efficient allocation of talents. Econ. Model. **24**(4), 648–672 (2007)
6. Eurostat Homepage: http://ec.europa.eu/eurostat. Last accessed 6 Apr 2018

Chapter 26
(Ir-)Rationality of Teams: A Process-Oriented Model of Team Cognition Emergence

Iris Lorscheid and Matthias Meyer

Abstract In today's competitive environment, companies rely increasingly on teams and their flexibility. While effectively working teams may accomplish great results, ineffective teams may fall short of their potential and can even be a risk for the organization. Little is known about the socio-cognitive processes of team decisions and particularly the emergence of knowledge from individual to team level. This study addresses this process by analyzing team cognition as an emergent property. The here presented research approach allows for a deeper analysis of the underlying processes. A laboratory experiment provides information about quantitative patterns of individual and team cognition. For the analysis of these patterns, we introduce the team cognition matrix. By applying this format to the results of the laboratory experiment, this study identifies four categories of typical emergent team cognition structures. These four categories are the basis for a simple decision algorithm that was analyzed in an agent-based model. The resulting simulation shows that 67% of all simulated group decisions are very close to the empirical group decisions (ranking position distances $\leq$3) and 89% are close on a medium range (ranking position distances $\leq$6). The article contributes to the current literature by showing an innovative research approach that further is applied to open up the black box of successful team behavior beyond well-known static attributes.

Keywords Team performance · Decision dynamics · Interactive team cognition · Laboratory experiments · Agent-based modeling

---

I. Lorscheid (✉)
University of Applied Sciences Europe, Hamburg, Germany
e-mail: iris.lorscheid@tuhh.de

M. Meyer
Hamburg University of Technology, Institute of Management Accounting and Simulation, Hamburg, Germany

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_26

26.1 Introduction

The cultural anthropologist Margaret Mead formulated "never doubt that a small group of thoughtful committed people can change the world. Indeed, it is the only thing that ever has" ([23], unit 3). The corporate world is indeed shaped by group decisions, such as by management boards or project teams. Teams are an important element in organizations in order to ensure the necessary flexibility and adaptability of organizational behavior facing competitive challenges [7]. There is common sense on the fact that groups can make better decisions than individuals, because groups may bring diverse information, skills, and perspectives together and at the same time reduce individual limitations through biases or personal preferences [8, 19]. Consequently, group decisions are a central element in organizational decision-making.

But groups do not always outperform individuals. Groupthink theory, for example, describes dysfunctional patterns of interaction that may result in poor group decisions [9, 11, 12]. Pressure toward uniformity, group homogeneity, or other social and contextual influences may suppress individual contributions or effective group decision-making [19]. Social and applied psychology literature formulates determinants of effective team decision-making with reference to social, behavioral, cognitive, and contextual factors ([17], p.271). The literature also discusses properties of effective teams. For example, Katzenbach and Smith [13] describe effective teams as teams with an appropriate number of members with diverse skills and viewpoints, clearly defined goals, explicit roles, and rules for members.

Interactive team cognition theory emphasizes team cognition as an activity, not a property, and describes teams as cognitive (dynamical) systems in which "cognition emerges through interaction" ([4], p.256). The call for an emphasis of research studies on the socio-cognitive interaction processes in teams is supported by many [6, 10, 18, 24]. The processes and mechanisms of how team cognition emerges from individuals are still a black box. However, these dynamic processes determine team performance. In line with this stream of research, this study contributes to the exploration of "how and why organizationally relevant outcomes emerge rather than focus only on differences in what has already emerged" ([10], p.45).

Further, this research relates to research on social ties and team reasoning in experimental economics and social psychology. This literature stream analyzes the effect of social connectedness in teams on coordination tasks [1–3]. Given the results from these studies, teams may improve their group performance by team-reasoning strategies or through perceived high values of social ties to the group members.

This study uses a laboratory experimental setup to explicate individual and team cognition structures and to analyze the processes triggered by these structures to identify dominant patterns of team cognition emergence. On this basis we use an agent-based model to model the processes how team cognition emerges from individual team cognition. The fit between the simulation results and the team results is analyzed and is used as an assessment whether the hypothesized dynamic processes are a possible explanation for observed team performance [14, 22].

26.2 Laboratory Experiment

The core of the laboratory setup is a team building game [15]. The task is to rank 15 listed items. The story behind this ranking task is a hypothetical accident at sea. Fire leads to an emergency in which the team has to leave a sinking ship. The 15 items on the list are possible things to carry from the sinking boat to a lifeboat. The task is to evaluate these items regarding their relevance for survival. A handout describes the story, together with the list and explanation of the pieces.

The groups consist of three, four, or five participants. In the first round, the participants are asked to rank all items individually on a ranking chart handout, without talking to each other. They have 10 minutes for this task and may not see the ranking results of others. Afterward, the group is asked to discuss and agree on one group ranking, which is noted on another ranking chart. This round has a time limit of 20 minutes.[1] The group discussions are audio recorded. These records provide information on the interaction dynamics in the group, such as conversational turn-taking or leadership. After the group discussion, the game is over,[2] and the individual and group rankings may be evaluated.

The experimental setup allows for the evaluation of the performance of the individual and team. A ranking of the US coast guard is the basis for this evaluation. This ranking provides an expert evaluation of the listed items with a description of the reasoning behind. Given this, we calculate individual and team scores.

The scores measure the individual and team performance by calculating how close the individual and team decisions match the expert opinion. The individual score is the sum of the absolute differences between the individual ranking positions and the expert ranking positions. For example, if the participant ranks one item to position 7 and the coast guard ranking has the same piece on ranking position 3, the absolute difference for this item is 4. The sum of this calculation for all 15 items results in the individual score. The group score is calculated accordingly by the sum of the absolute differences between the group ranking positions and the expert ranking. The smaller the score, the better the performance.

These scores allow for a comparison of individual and team performance regarding this task. For the quantification of this comparative evaluation, we introduce the performance measures *(P1) team result*, *(P2) team intelligence*, and *(P3) team stupidity* (see Table 26.1). These measures were adapted and extended from Cooke and Kernaghan [5].

The performance measures evaluate the group result with regard to the average results of the team members (P1), the best team member result (P2), and the worse team member result (P3). The average of the individual scores provides a benchmark to analyze whether the group interactions result in a gain or a loss in comparison to

[1] The time limits for the game rounds are set according to the team building game.

[2] Further adaptations could include another individual ranking after the group discussion to analyze the change of individual decisions and explicate the hidden knowledge that was not part of the group ranking.

Table 26.1 Team performance – performance measures

Performance measure	Calculation	Interpretation
(P1) Team result	*team_score – median(individual_scores)*	<0 → *gain* >0 → *loss*
(P2) Team intelligence	*team_score – min(individual_scores)*	<0 → *team intelligence*
(P3) Team stupidity	*team_score – max(individual_scores)*	>0 → *team stupidity*

the individual results. If the group score is lower than the best individual score, the team succeeded in achieving a better result than all group member alone. In this case we identify a case of team intelligence. If the group score is higher than the highest individual score, the group performed worse than all group members, and we identify a case of team stupidity.

Looking at the results of the laboratory experiments (see Fig. 26.1), we identify varying results regarding team performance. 23 of 26 groups gain through the group interaction (see P1). The groups have a better score than the average of all individual scores. We identify team intelligence (P2) for nine groups. Here, the group performed better than the best performing individual. In three of these groups, we even see big jumps, from −10 up to −18 in comparison to the best individual score (see groups X, A, and I). The three worst performing groups in the data set have a loss in team performance, meaning that the team score was worse than the average score of all individuals. For group T the group score even matches the weakest individual score.

These findings raise questions, such as the following: What makes one group perform better than others? What were the interaction processes that lead to the emergent phenomena of team intelligence? What hinders less successful groups in the decision process to achieve the better results of strong groups?

The setup of the laboratory experiment is a step forward opening the black box of the decision process. The available data allow for a structured analysis of the ranking patterns behind the aggregated evaluation measures. The individual and group rankings explicate decisions on both levels and provide the basis for the analysis of the emergent group decision dynamics. In the next section, we introduce the team cognition matrix as format to analyze the processes behind the performance.

26.3 Team Cognition Matrix

The ranking patterns of the laboratory experiment explicate the individual and team cognition structures underlying the group decision-making process. This structure provides a data basis for the analysis of the emergent process from many perspectives, such as the relative ranking positions in a group or the distances of ranking positions per item. For the evaluation of the team structures for each of the 15 ranking items, we developed the *team cognition matrix* (see Fig. 26.2).

groups					team performance		
groupID	group size	median of ind. scores	min ind. score	max ind. score	(P1) team result	(P2) team intelligence	(P3) team stupidity
X	5	60	48	68	-30	-18	-38
A	4	53	38	58	-25	-10	-30
Z	5	64	34	94	-22	8	-52
I	5	60	58	90	-20	-18	-50
Y	5	56	34	86	-20	2	-50
S	3	56	52	64	-18	-14	-26
K	4	69	56	74	-17	-4	-22
M	4	57	28	96	-17	12	-56
Q	4	67	54	102	-17	-4	-52
E	4	64	36	74	-16	12	-26
N	4	61	44	82	-15	2	-36
W	4	59	34	72	-13	12	-26
D	4	70	44	82	-10	16	-22
L	4	62	58	66	-10	-6	-14
C	4	55	36	70	-9	10	-24
J	4	51	42	68	-9	0	-26
B	4	54	44	72	-8	2	-26
F	4	73	64	82	-7	2	-16
H	4	78	48	80	-6	24	-8
P	4	52	50	56	-6	-4	-10
O	4	46	38	50	-4	4	-8
R	3	46	44	52	-4	-2	-10
U	5	58	40	74	-2	16	-18
G	5	56	46	76	0	10	-20
V	4	68	42	78	0	26	-10
T	5	68	38	80	12	42	0

Fig. 26.1 Team performance results ($N = 109$ participants, 26 groups, sorted by (P1) team result)

individual distances | individual distances to group | individual distances to benchmark

$$\begin{bmatrix} 0 & a_{1,2} & a_{1,3} & \cdots & a_{1,g} & a_{1,g+1} & a_{1,g+2} \\ a_{2,1} & 0 & a_{2,3} & \cdots & a_{2,g} & a_{2,g+1} & a_{2,g+2} \\ a_{3,1} & a_{3,2} & 0 & \cdots & a_{3,g} & a_{3,g+1} & a_{3,g+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{g,1} & a_{g,2} & a_{g,3} & \cdots & 0 & a_{g,g+1} & a_{g,g+2} \\ a_{g+1,1} & a_{g+1,2} & a_{g+1,3} & \cdots & a_{g+1,g} & 0 & a_{g+1,g+2} \\ a_{g+2,1} & a_{g+2,2} & a_{g+2,3} & \cdots & a_{g+2,g} & a_{g+2,g+1} & 0 \end{bmatrix}$$

Fig. 26.2 Team cognition matrix (TCM) – overview of the cognitive structure of one group with group size g, filled for one ranking item

This matrix gives an overview of the group characteristic per ranking item in a condensed manner. The matrix is quadratic, with n rows and n columns. The order n of the matrix depends on group size g. The matrix elements are absolute distances

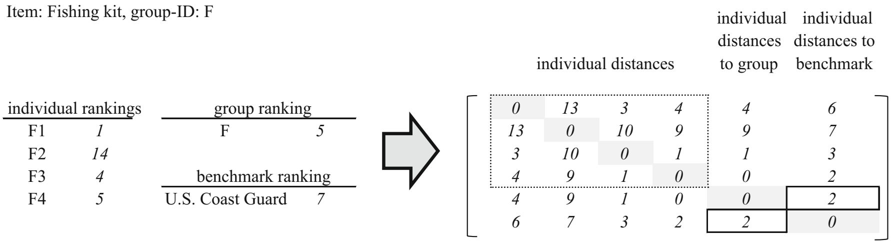


Fig. 26.3 Example – ranking results and the resultant team cognition matrix for the item "fishing kit" and group-ID "F," with four team members

between ranking positions. Rows and columns with indices from 1,1 to g,g contain distances between two individual ranking positions for all group members and all possible pairwise combinations. The elements in the row and column $g + 1$ show the distances between the respective individual ranking position and the group decision and the elements in row and column $g + 2$ the distances to the benchmark ranking, here defined by the expert ranking. Because the direction of deviation is not in the focus, but only the information of how far away the positions are from each other, the matrix contains absolute and unsigned distances. The main diagonal has zero values as these elements are self-referential. The entries in the matrix are filled symmetrically identical over the main diagonal, to simplify the analysis process.

Figure 26.3 shows an example of the calculation of the team cognition matrix. The left hand of the figure shows the individual rankings and group ranking for the item "fishing kit" and the group with group-ID "F," as given from the laboratory experiment data set. The experimental setup specifies 15 ranking items. Consequently, the range of possible distance values in the matrix is the interval of [1,14]. The group in this example has four members. Group member F1 ranked the item on position 1. This ranking position leads to the following distances to the other group members: 13 to group member F2, 3 to group member F3, and 4 to group member F4. These values are the first four elements of the first row in the team cognition matrix (the right part of Fig. 26.3). The next value of this row shows the distance of F1 to the group decision, being four positions away from the group ranking position 5. The expert opinion on the relative relevance of the fishing kit is ranking position 7. The last value of the first row shows the distance of the individual ranking by F1 to this benchmark ranking, being six positions away from each other. The matrix contains all distances for all group members accordingly. Additionally, the element on a $g + 1, g + 2$ shows the distance between the group ranking position (here: 5) and the benchmark ranking position (here: 7).

Looking through this exemplary team cognition matrix, we identify two individuals being very close to each other in their ranking positions (F3 and F4), while another team member is far away to the other individual ranking positions (F2). The fourth individual F1 has a small- to medium-range distance to F3 and F4. Given these distances, we can identify a general heterogeneity of individual ranking positions for the fishing kit in this group. At the same time, there is a majority with

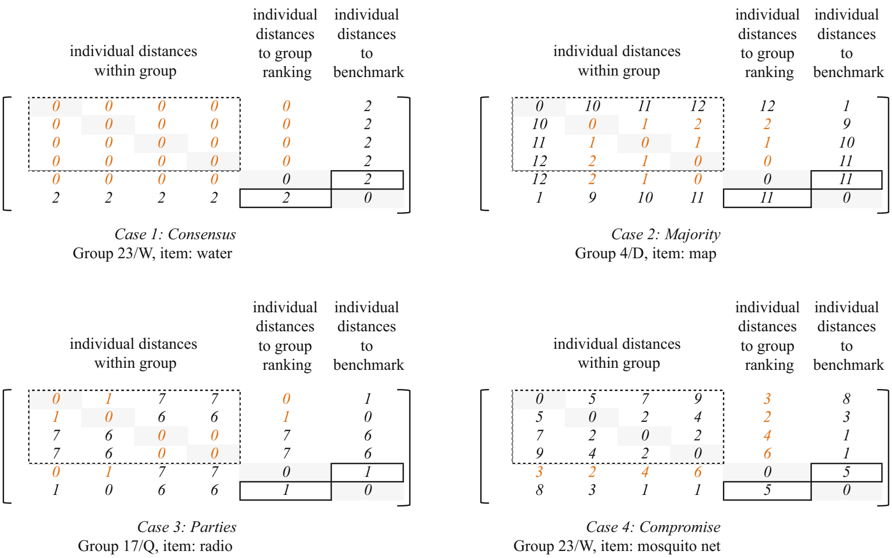


Fig. 26.4 Team cognition structure – identified categories

very close ranking positions of 4 and 5, and the group agreed on ranking position 5. This indicates a possibility that a majority decision to this ranking position took place. The majority and group decision was also closer to the benchmark than the other team member ranking positions. Thus, the better decision prevails in this example.

The matrix structure allows for such an evaluation of emergent team results by analyzing and identifying individual and team cognition structure characteristics, their frequency of occurrence, and their explanatory power.

26.4 Results

For the analysis, we created team cognition matrices[3] for each group and item. The set of matrices provides information on the distances among group member rankings (representing the homogeneity of individual cognitive structures), the distances between individual and group rankings (representing the degree of consensus between individual and group rankings), and the distances to the benchmark (representing the individual and team performances). On closer inspection, the matrices reveal four typical and reoccurring cognition structures (Fig. 26.4).

[3] In total 390 matrices for 15 items and 26 groups.

The individual distances among group member rankings (the areas within the dotted lines) represent the divergence of positions regarding the item in the group. In the example of case 1 (consensus), all members of group 23 (W) individually ranked "water" as the most relevant item for survival. Consequently, all values are zero; the cognitive structure among the team members is homogeneous.

The distances between individual rankings and group decisions reveal the degree of consensus between the individual rankings and group ranking. In case 1 these values are zero for all team members. The group agreed on the most relevant item and ranked water on the top. We often identify such consensus decisions for homogenous individual team member rankings. Case 1 shows an example of this situation. It is easier to come to an agreement when individual positions are close to each other or even equal, which is reflected in this example. We observe this situation frequently for the item "water" in the laboratory experiment. Given the qualitative data from observations and audio records, the decision concerning this item is usually the first and easiest agreement in the group discussion. Regarding the performance, this decision reached individual and team scores of 2, as the expert opinion ranked water on position 3.

More complicated situations occur if positions are far apart from each other. Case 2 (majority) shows the team cognition structure for group 4 (D) and the item "map." As one can see, there is a majority of 3 team members with close ranking positions and 1 team member with high distances being 10–12 positions away from the group. Looking at the distances between individual and group ranking, we see that the majority determined the group decision. We find this majority structure frequently. Often, but now always, the majority defines the group ranking position.

Interestingly, the majority of group 4 (D) was not right in comparison to the expert opinion for the item "map." We identify a team score of 11. The individual with the single opinion, however, was very close to the benchmark with an almost perfect individual score of 1. We find both right and wrong majority decisions in the data set.

Case 3 (parties) shows another typical team cognition structure. The team members split into two parties. Here, two parties of two team members have close ranking positions, but the two parties are far away from each other (by 6 and 7 ranking positions). Group 17 (Q) ranked the item "radio" close to one party. This is a frequently observed structure for this case. One party gets its way and defines the group decision. In this example, the winning party of group 17 was also the party with the better scores. In our data set, we again identify both, winning successful and winning unsuccessful parties.

Finally, we identify situations in which the individual ranking positions scatter, so that no majority and no party of ratings establish. If no majority evolves, groups find a compromise between individual ranking positions. Case 4 (compromise) shows such an example for group 23 (W) and the item "mosquito net."

To further explicate the possible connection between team cognition structure and performance, we implement the identified patterns as decision dynamics in a simulation study. By this we aim to test the explanatory power of these structures for the emergent group outcome and performance. For this, we apply agent-based

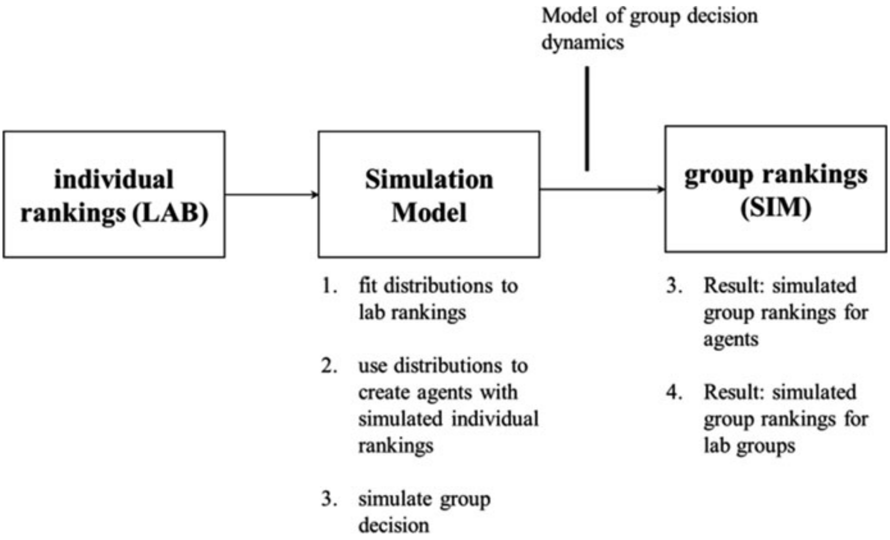


Fig. 26.5 Simulation model analysis of group decision dynamics

simulation [16, 20, 21] and endow the agents with group decision dynamics that use the team cognition structures as criteria for the group decision-making and agreement on group ranking items. For the implementation of the simulation model and the analysis of the empirical and simulated data, we use the programming language R. The data analysis of the laboratory experiment and the simulation output analysis run through the same code lines in R. This reduces errors and facilitates the validation process.

Figure 26.5 provides an input-output diagram and description of the main process elements of the simulation study. The individual rankings from the laboratory experiment serve as model input. Given these rankings, the simulation model fits statistical distributions to the input data. In the result of this first step, we have a specification of distributions and distribution parameters for each item that best represent the empirical distribution of individual rankings.

In the second step, the model uses the distributions to create artificial agents with simulated individual rankings. For this, the algorithm chooses random ranking positions for each item out of the statistical distributions. To avoid duplicates or not assigned ranking positions, the algorithm sorts the resulting random values and assigns the ranking positions 1–15 to this sorted list of items. In case of equal random numbers, one is randomly chosen first in the simulated individual ranking. As a result of this step, the simulation model creates agents with individual rankings with characteristics according to the empirical data of the laboratory experiment.

After the individual rankings are generated, the agents are assigned to teams of four and enter the group decision process. The group decision dynamic has three phases: (phase 1) identification of the team cognition structure per item, (phase 2) decision for a group ranking position per item, and (phase 3) ranking refinement per group. The model runs through all these phases for every group.

The parameter *consensLimit* supports the identification of the team cognition structure category. A distance between two individual ranking positions has to be smaller or equal than *consensLimit*, so that this distance is evaluated as “close enough” for evaluating consensus between two positions. This comparison of distances to *consensLimit* is the core decision for the structural evaluation of each team cognition matrix in the simulated group decision. In the current simulation experiments, we set the value of *consensLimit* to 3, which will be further analyzed in robustness tests.

In phase 1 the algorithm evaluates if the cognitive structure represents the category *consensus*, *majority*, *party*, or *none* of these three. If all individual ranking position distances are smaller than *consensLimit*, all group members have close ranking positions, and the program identifies *consensus* in the group. If this is not true, the algorithm continues and checks if a majority of group members has close ranking positions below this limit. If so, the algorithm identifies a *majority* structure for this item. Otherwise, if no majority structure was identified, the program proceeds to check if the group splits into two parties of ranking positions. In this case, the model evaluates a *party* structure. If the simulation does not find a party structure, none of the three chosen categories matched and the cognitive structure for the item is assigned to the fourth type labeled with *others*.

The identified category of phase 1 determines the group ranking position for each item in phase 2. If the group shows *consensus* in the individual ranking positions, the group agrees on one randomly chosen[4] individual ranking position within the consensus range. Otherwise, if the group has a *majority*, one of the individual decisions out of the majority determines the group decision. If there is no majority but a *party* structure in the group for an item, one randomly chosen party wins, and one of the individual rankings in the winning party defines the group ranking position. If the cognitive structure belongs to *others*, the group chooses a *compromise*, being the average of all individual ranking positions for this item.

The resulting group ranking positions from phase 2 need to be refined in the third phase of the group decision dynamic. Phase 2 may potentially result in duplicates as well as in not assigned ranking positions. However, each of the 15 ranking positions should only occur once, without leaving one ranking position out. For keeping this requirement, the ranking positions are sorted and re-ranked in the same logic as the simulated individual rankings are (see before). This refinement process may lead to small deviations from the group decision dynamics, but keeps the overall, relative relevance of items.

This group decision dynamics of the simulation model can be applied in two ways. In one type of application, the simulation model only considers the groups from the laboratory experiment and their individual empirical rankings.

[4] In further research, we aim to consider alternative selection strategies, such as choosing the individual ranking position that appears most frequently within the consensus range.

The algorithm simulates the group decisions, which can be then compared to the data from the lab. In the other type of application, the simulation model simulates the group decisions for the artificial agents and their rankings. As a result, we may derive more general knowledge about the characteristics of the emergent team cognition structures under a wide variety of individual structures and combinations.

We apply the decision algorithm to simulate the group decisions in the laboratory experiments. Figure 26.6 shows the empirical distribution of group ranking positions. Figure 26.7 shows the distribution of the simulated group ranking positions. The algorithm uses the individual rankings and groups from the laboratory experiment and applies the group decision dynamics as implemented to calculate the group rankings. The violin plots provide an aggregated view on the resulting distributions over all groups. Looking at the distributions, we identify for most items the same tendencies regarding high, medium, or low median values over the elements. The item "mosquito-net" (mosq_net), for example, has higher-ranking positions in both results, while the item "nylon-rope" (nylon_r) has distributions around medium-range ranking positions in both figures. However, we also identify broader distributions in the simulated data set than in the empirical data set.

For evaluating the fit and explanatory power of the decision dynamics, we go beyond the aggregated distributions and compare every empirical and simulated ranking position for each group and item (see Fig. 26.8). For this analysis, we calculate the absolute distances between the result of the laboratory experiment and the simulation output. Having these distances, we create limits of distances (see the first column of Fig. 26.8) and count the number of observations per item in which we identify distances within these limits.

Finally, we express the number of observations relative to the total number of observations per item (26 observations per item, 1 observation per group). Thus, the value 0.12 for sextant and distances of 0 means that in 12% of 26 group rankings, the simulated ranking position for sextant matches precisely the empirical result. The last column (all items) expresses the relative number of observations with distances

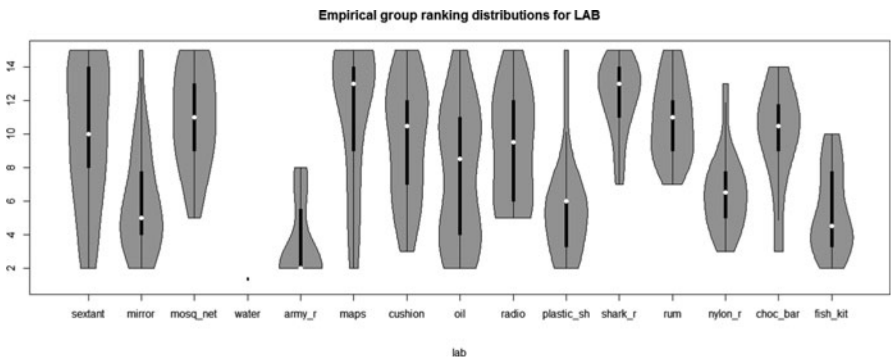


Fig. 26.6 Output distribution of empirical group rankings – empirical data from laboratory experiment (26 groups)

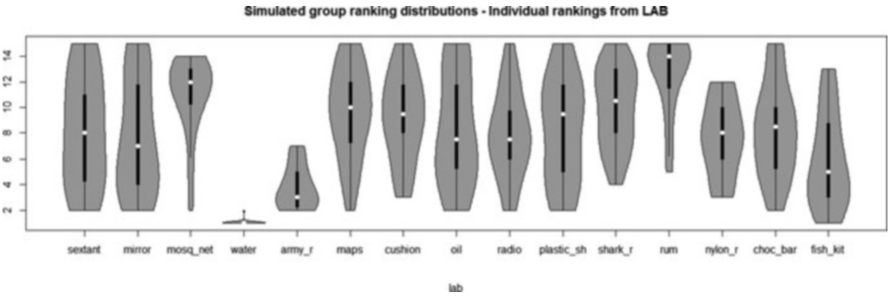


Fig. 26.7 Output distribution of simulated group rankings – simulated group decisions with empirical individual rankings from laboratory experiment (26 groups)

	relative number of observations															
distances (empirical ./. simulated group ranking)	sextant	mirror	mosq_net	water	army_r	maps	cushion	oil	radio	plastic_sh	shark_r	rum	nylon_r	choc_bar	fish_kit	***all items***
0	0.12	0.04	0.12	0.96	0.35	0.08	0.04	0.19	0.15	0.04	0.15	0.08	0.04	0.04	0.12	**0.17**
≤ 3	0.54	0.42	0.62	1.00	0.92	0.54	0.65	0.65	0.85	0.58	0.62	0.62	0.65	0.58	0.77	**0.67**
≤ 6	0.77	0.77	0.92	1.00	1.00	0.92	0.92	0.88	0.92	0.77	0.88	0.88	0.92	0.81	0.96	**0.89**
≤ 9	0.96	0.88	0.96	1.00	1.00	0.96	1.00	0.96	1.00	0.96	1.00	1.00	1.00	0.96	0.96	**0.97**
≤ 12	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.00**
≤ 14	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**1.00**

Fig. 26.8 Evaluation of fit between empirical and simulated group rankings – entries: relative number of observations with distances within the limits (cumulative). Distances are calculated between empirical and simulated ranking positions for all groups and items (26 groups)

within limits. Thus, in 17% of all observations (390 for 26 groups and 15 items), the simulation output matches the empirical result exactly, in 67% of all observations we find a distance of smaller or equal 3, and so on. The quality of the simulated group rankings is in particular remarkable for the items with an overall wide distribution of ranking positions. For the item "oil," for example, we identify all possible ranking positions between 1 and 15 in the empirical data set. Nevertheless, the decision algorithm replicates the exact ranking position in 19%, a close ranking position in 65% of all cases, meaning in absolute values that in 5 out of 26 observations the algorithm was exact and in 17 out of 26 observations the algorithm was successful. Overall, this result improves for items with more narrow empirical distributions.

The simulation results show that the simple decision dynamics can already explain 67% of all group decisions. The results have to be revised with more data from further laboratory experiments and more simulation runs. With more observations, we aim to identify further cognition structures and refined group processes that improve the explanation of the emergent team cognition and, in the end, team performance.

26.5 Conclusion

Little is known about the socio-cognitive processes of team decisions and particularly the emergence of knowledge from individual to the team level. This study addresses this process by investigating team cognition as an emergent property. We introduced the team cognition matrix that explicates the underlying cognitive structure in groups. By applying this format, we identified four categories of group decision dynamics. These categories were the basis for a decision algorithm that we implemented in an agent-based simulation model. Our approach shows a way of how to find possible explanations for the "how and why organizationally relevant outcomes emerge" ([10], p.45) and identifies underlying patterns. The good fit between the simulation results and the empirical results supports the hypothesis that the modeled dynamic processes partly explain the observed team performance. Future research has to test and to refine the hypotheses about the four identified constellations and the decision processes triggered by them. Also, alternative or refined process models have to be developed and tested. In further research with additional data sets and more analyses of resulting team cognition matrices, we hope to identify more details concerning team cognition patterns and micro-level processes of relevance.

References

1. Attanasi, G., Hopfensitz, A., Lorini, E., Moisan, F.: Social connectedness improves co-ordination on individually costly, efficient outcomes. Eur. Econ. Rev. **90**, 86–106 (2016)
2. Bacharach, M.: Beyond Individual Choice: Teams and Frames in Game Theory. Princeton University Press, Princeton (2006)
3. Colman, A.M., Pulford, B.D., Rose, J.: Collective rationality in interactive decisions: evidence for team reasoning. Acta Psychol. **128**(2), 387–397 (2008)
4. Cooke, N.J., Gorman, J.C., Myers, C.W., Duran, J.L.: Interactive team cognition. Cogn. Sci. **37**(2), 255–285 (2013)
5. Cooke, R.A., Kernaghan, J.A.: Estimating the difference between group versus individual performance on problem-solving tasks. Group Organ. Stud. **12**(3), 319–342 (1987)
6. Crusius, J., van Horen, F., Mussweiler, T.: Why process matters: a social cognition perspective on economic behavior. J. Econ. Psychol. **33**(3), 677–685 (2012)
7. Delgado, P.M.I., Martínez, A.M.R., Martínez, L.G.: Teams in organizations: a review on team effectiveness. Team Perform. Manag. Int. J. **14**(1/2), 7–21 (2008). https://doi.org/10.1108/13527590810860177
8. DeVilliers, R., Woodside, A.G., Marshall, R.: Making tough decisions competently: assessing the value of product portfolio planning methods, devil's advocacy, group discussion, weighting priorities, and evidenced-based information. J. Bus. Res. **69**(8), 2849–2862 (2016). https://doi.org/10.1016/j.jbusres.2015.12.054
9. Esser, J.K.: Alive and well after 25 years: a review of groupthink research. Organ. Behav. Hum. Decis. Process. **73**(2), 116–141 (1998). https://doi.org/10.1006/obhd.1998.2758
10. Grand, J.A., Braun, M.T., Kuljanin, G., Kozlowski, S.W., Chao, G.T.: The dynamics of team cognition: a process-oriented theory of knowledge emergence in teams. J. Appl. Psychol. **101**(10), 1353 (2016)

11. Janis, I.L.: Victims of Groupthink: a Psychological Study of Foreign-Policy Decisions and Fiascoes. Houghton-Mifflin, Boston (1972)
12. Janis, I.L.: Groupthink. Houghton-Mifflin, Boston (1982)
13. Katzenbach, J.R., Smith, D.K.: The Wisdom of Teams: Creating the High-Performance Organization. Harvard Business Review Press, Boston (2015)
14. Klingert, F.M.A., Meyer, M.: Effectively combining experimental economics and multi-agent simulation: suggestions for a procedural integration with an example from prediction markets research. Comput. Math. Organ. Theory. **18**(1), 1–28 (2012)
15. Knox, G.: Lost-at-sea – team building game (2016). Retrieved 28 Nov 2016, from http://insight.typepad.co.uk/insight/2009/02/lost-at-sea-a-team-building-game.html
16. Macal, C.M.: Everything you need to know about agent-based modelling and simulation. J. Simul. **10**(2), 144–156 (2016)
17. Reader, T.W.: Team decision making. In: The Wiley Blackwell Handbook of the Psychology of Team Working and Collaborative Processes, pp. 271–296. John Wiley & Sons, Ltd, West Sussex (2017)
18. Rousseau, V., Aubé, C., Savoie, A.: Teamwork behaviors: a review and an integration of frameworks. Small Group Res. **37**(5), 540–570 (2006)
19. Schulz-Hardt, S., Jochims, M., Frey, D.: Productive conflict in group decision making: genuine and contrived dissent as strategies to counteract biased information seeking. Organ. Behav. Hum. Decis. Process. **88**(2), 563–586 (2002). https://doi.org/10.1016/S0749-5978(02)00001-8
20. Secchi, D.: A case for agent-based models in organizational behavior and team research. Team Perform. Manag. Int. J. **21**(1/2), 37–50 (2015). https://doi.org/10.1108/TPM-12-2014-0063
21. Smaldino, P.E., Calanchini, J., Pickett, C.L.: Theory development with agent-based models. Organ. Psychol. Rev. **5**(4), 300–317 (2015). https://doi.org/10.1177/2041386614546944
22. Smith, E.B., Rand, W.: Simulating macro-level effects from micro-level observations. Manag. Sci. **64**, 5405–5421 (2017)
23. Verderber, K.S., Verderber, R.F., Sellnow, D.D.: Communicate!. Cengage Learning, Boston, MA (2013)
24. Woolley, A.W., Chabris, C.F., Pentland, A., Hashmi, N., Malone, T.W.: Evidence for a collective intelligence factor in the performance of human groups. Science. **330**(6004), 686–688 (2010)

Chapter 27
Early Holocene Socio-Ecological Dynamics in the Iberian Peninsula: A Network Approach

Sergi Lozano, Luce Prignano, Magdalena Gómez-Puche, and Javier Fernández-López de Pablo

Abstract Late Glacial and Early Holocene environmental changes affected different domains of human demography, settlement, and subsistence patterns. The variable spatial patterning produced by the prehistoric hunter-gatherers' archaeological record demands new approaches for analysing the multi-scalar nature of human-environmental interactions. In this contribution, we presented part of a long-term research programme aimed to cover this gap in the context the Iberian Peninsula from the Late Magdalenian to the end of the Late Mesolithic.

PALEODEM ("Late Glacial and Postglacial Population History and Cultural Transmission in Iberia (c.15,000–8,000 cal BP)") is a Consolidator Grant ERC project that addresses the role of human population levels and geographical distribution over the relationship between climatic events and cultural dynamics in our context of study. To do so, it will develop a three-level (micro-regional, regional, and macro-regional) analysis. In this contribution, we will focus on the macro-regional scale, which is addressed through a combination of network analysis and computational modelling.

Keywords Archaeology · Paleodemography · Networks analysis · Modelling of dynamics on and of networks · Iberian Peninsula

S. Lozano
Departament d'Història Econòmica, Institucions, Política i Economia Mundial, Universitat de Barcelona, Barcelona, Spain
e-mail: slozano@iphes.cat

L. Prignano
Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona, Spain

Universitat de Barcelona Institute of Complex Systems (UBICS), Universitat de Barcelona, Barcelona, Spain

M. Gómez-Puche · J. F.-L. de Pablo (✉)
I.U. de Investigación en Arqueología y Patrimonio Histórico (INAPH), University of Alicante, San Vicente del Raspeig, Spain
e-mail: jfernandez@iphes.cat

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_27

27.1 Introduction

Late Glacial and Early Holocene environmental changes affected different domains of human demography, settlement, and subsistence patterns. The variable spatial patterning produced by the prehistoric hunter-gatherers' archaeological record, from local bands to larger regional groups, demands new approaches for analysing the multi-scalar nature of human-environmental interactions. In this contribution, we present different aspects of a long-term research programme aimed to decipher the relationship between demographic dynamics and cultural transmission in the Iberian Peninsula from the Late Magdalenian to the end of the Late Mesolithic.

The contribution had two parts. First, we briefly introduced PALEODEM ("Late Glacial and Postglacial Population History and Cultural Transmission in Iberia (c.15,000–8,000 cal BP)"), a Consolidator Grant recently awarded by the European Research Council (ERC) to support part of this research programme. PALEODEM addresses the role of human population levels and geographical distribution over the relationship between climatic events and cultural dynamics in our context of study. To do so, it develops a three-level (local, regional, and macro-regional) analysis. First, at a local scale, PALEODEM is studying the impact of climate change and hydrological stress on different human settlement areas. Second, it is reconstructing population patterns at four different Iberian regional units analysing summed probabilities of radiocarbon dates. Finally, it will conduct network analysis and modelling at a macro-regional scale to examine longitudinal variations in settlement networks affecting the overall connectivity amongst hunter-gatherer populations.

The largest part of our contribution to this session focused on PALEODEM's macro-regional approach, which studies how changes in population levels and the spatial network linking human communities across regions could have influenced different cultural macro-phenomena.

In particular, we address the following two interrelated questions:

1. What was the role played by human groups in the diffusion of innovations across regional units, depending on their population, geographic location, and position within the network structure?
2. To what extent the changes in the spatial configuration and density of social networks could influence macro-phenomena like the rapid diffusion of certain technological innovations or the loss of technical skills?

27.2 Methodology

We aim at answering these questions by analysing two cultural macro-phenomena previously identified in the archaeological record of our context of study: the loss of technical skills during the Epipalaeolithic and the rapid diffusion of compound projectile points with trapeze shape during the Late Mesolithic.

In order to study these phenomena, we have designed a methodology based on the use of regional population estimates (previously generated as part of PALEODEM's efforts) and a methodology combining network analysis and multi-agent computational modelling. Specifically, for each one of the two phenomena, we will:

1. Construct spatial networks corresponding to the scenario before and after the corresponding temporal windows, where nodes represent archaeological sites (i.e. human groups) and links are defined from chronological overlapping, spatial proximity, and/or cultural attributes [1, 2].
2. Compare the two snapshots based on different structural characteristics [3]. In particular, such a study will include both properties addressing networks as a whole (e.g. degree distributions, assortativity, centralities, and brokerage roles), as well as local features (e.g. clustering coefficient and motifs' identification).
3. Conduct experiments to model dynamics on and of networks [4]. If the structural comparison in the previous step does not reveal significant differences, our model will consider the network as a static substrate supporting cultural transmission processes (i.e. basically, adoption of cultural traits) among agents (human groups). Thus, our goal in this case would be assessing how the structural aspects of this static network could have influenced the macroscopic cultural phenomenon under study (i.e. modelling on static networks). On the contrary, if the network comparison in the second step uncovers significant differences between the two snapshots, we will consider a more complex (and realistic) modelling scenario where both cultural attributes and structural features evolve along time (modelling on dynamic networks).

27.3 Expected Impact

By developing the proposed research programme, we expect our work to contribute to the integration of network analysis and computational modelling of cultural dynamics into a fast-growing literature on quantitative studies of demographic processes in pre- and proto-history[1]. More generally, this work aims at contributing to the consolidation of a rather unexplored application of network science to the study of prehistoric long-term cultural dynamics.

Acknowledgements This research was funded by the European Research Council Consolidator Grant-2015 number 683018 to JFL for the PALEODEM Project under the European Union's Horizon 2020 research and innovation programme. Authors also acknowledge financial support by the Generalitat de Catalunya through the SGR programme (Projects No. 2017SGR-836) and the Generalitat Valenciana for the CIDEGENT/2018/040 grant.

[1]For some context on this issue, check https://crossdem18.wordpress.com/page/ and https://crossdem2019.com/.

References

1. Coward, F.: Grounding the net: social networks, material culture, and geography in the Epipalaeolithic and early Neolithic of the Near East. In: Knappett, C. (ed.) Network Analysis in Archaeology: New Approaches to Regional Integration. OUP, Oxford (2013)
2. Prignano, L., Morer, I., Diaz-Guilera, A.: Wiring the past: a network science perspective on the challenge of archeological similarity networks. Front. Digit. Humanit. **4**, 13 (2017)
3. Newman, M.E.J.: Networks: An introduction. Oxford University Press, Oxford, UK (2010)
4. Lozano, S.: Dynamics of social complex networks: some insights in recent research. In: Dynamics On and Of Complex Networks: Applications to Biology, Computer Science and the Social Sciences. Series, Modelling and Simulation in Science, Engineering and Technology. Springer-Birkhauser, Boston (2009)

Chapter 28
Influences of Innovation in Market Value

Nuno Trindade Magessi and Luis Antunes

Abstract Innovation has a great and sometimes decisive influence on companies' competitive advantage. However, innovations based on artificial intelligence such as robotics or computerisation can destroy jobs. All jobs are at risk. In this sense the lack of jobs will affect the market value of the companies that heavily invested in innovation. Based on this preoccupation, a multi-agent model denominated INOV-I has been developed to replicate this tendency in labour markets. We ran the model with three issues into consideration: (a) the proportion of innovative companies present in an artificial economy, (b) speed-to-market and time to delivery, and (c) the mobility effect. Several simulations yielded preliminary results that demonstrated that innovation has indirect negative impacts on both real and potential market value.

Keywords Robotics innovation · Market value · Speed-to-market effect · Workers' mobility effect · Labour market

28.1 Introduction

One of the issues behind innovation is the dissemination of value and the knowledge of the working force in developed countries. Progress in the sector of artificial intelligence has also been generating unemployment [1]. The alarms have started to ring, in response to recent innovations such as a machine capable of replacing most the accountants' tasks [2] or the case of the Tesla car factory which lacks any human workers in production. These facts and other cases like artificial managers or lawyers are transforming our societies and putting in risk people's value and the knowledge acquired all over the years. There are additional risks to countries' social security systems. The value generated by innovation is becoming an issue for not only lower-skilled but also more technically advanced workers, many of whom

N. T. Magessi (✉) · L. Antunes (✉)
Universidade de Lisboa, Faculdade de Ciências/BioISI – Biosystems and Integrative Sciences Institute, Lisbon, Portugal
e-mail: nmagessi@fc.ul.pt; xarax@ciencias.ulisboa.pt

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_28

have invested money into their knowledge. The greatest challenge now faced by top executive managers is what impact innovation will have on the value of the markets. The reasoning behind it is simple: if the current innovation processes either do not contemplate people or even replace them, then the reverse effect will take place for their companies, and they will naturally lose customers directly or indirectly. Clearly at this moment, we are facing a stand-off between the market value created by innovations and the potential risk to decrease its value represented by these same innovations [3].

This scenario leads us to ask if there is an ideal speed for implementing innovation. Is it possible to achieve a balance between AI innovations and market value? This article analyses the potential effect of this issue on the future of millions of workers all over the world. The risk exists that current-day efforts of the business sector to achieve efficiency through AI may end up destroying the very markets they want to achieve and grow.

To this end, it was designed a simulator under multi-agent-based systems methodology to assess the impact of the speed of innovation on market value.

This article has the following structure: the next section and Sect. 27.3 revise recent literature dealing with the effects of innovation velocity on market value. Section 27.4 looks at the multi-agent model developed to tackle this issue. Section 27.5 shows and discusses the results and then. Finally, Sect. 27.6 spells out our conclusions and suggestions for introducing future research.

28.2 Critical Advances in Innovation

A review of the literature regarding critical advances in innovation literature has pointed that innovation planning is often useful because it enables researchers to broaden their understanding of previously understudied areas of knowledge.

This type of research, however, is not particularly effective in detecting novel opportunities for innovation [4]. Executives tend to allocate innovation resources in a biased way, towards strengthening pre-existing business models. They can be overly focused on promoting innovation mainly regarding production. However, several researchers have identified that truly innovative breakthroughs arise when business explores new concepts and ideas rather than solutions from based on past experiences [4, 5]. They argue for the importance of healthy internal competition within innovative organisations, as do [6]. This competition breaks through inertia in order to find outlier ideas that might "to open up a new direction" [5]. Businesses are often observed spending many of their resources developing innovative processes or products to add value to the clients and, of course, to improve profit margins. This type of innovations is normally expensive and time-consuming and requires significant upfront investment, which means that "future returns on these investments are always uncertain" [5]. The researchers argue that "In the operations area, much of the innovations and cost savings that could be achieved have already been achieved". This is not true, because several improvements in production performance are currently underway that lead to cost reduction. Artificial Intelligence promises to be the ultimate frontier of substantive gains and increased efficiency and effectiveness (e.g. quality) within organisations. This represents a high-impact transformation with great impact, in terms of improvements on product quality, but also in scaling production and even in the innovation of business models. The latter is relevant because it potentially explores an unused source of future value. It is also more difficult for competitors to imitate and replicate this innovation, as compared to that of products or processes. Business model innovations may lead to sustainable performance advantages. Finally, there is the probability of creating a very powerful competitive tool that can be used to mitigate the risks of competitiveness and control the forces of competitors on market [5].

28.3 Innovation Speed and the (R)evolution in Labour Markets

Nowadays managers are highly aware of the need for speedy innovation [7] due to great competition to carry innovations to market as quickly as possible. The authors conducted a survey on the state of innovation worldwide and concluded that “overly long development times” were seen as the principal limitation to obtaining return on innovation and product development. This is an obstacle with capacity to compromise all the companies’ efforts in innovation and respective success. So, there is a special focus on developing new products and services quickly and delivering them to market as soon as possible [7]. They defend that although size can provide scale, constant innovation yields sustainable market competitiveness. Unfortunately, this research only focuses on the advantages for companies, while neglecting the collateral effects, in terms of innovative automatic processes and the labour market.

However, and in addition to the relevance of speed-to-market, time-based strategies are implemented nowadays in order to come up with important tools for obtaining competitive advantage [8]. The fact is that being faster sometimes is not always the best option [8]. Speed-to-market is also important due to fast-changing technology and meeting customer needs. It is normally positively associated with the success of new products success but market risks control the direct effect [8]. Instead it might be less relevant to the new product success when market conditions point to low market uncertainty. The authors’ results reveal that technological uncertainty does not influence the speed-success relationship. Innovation strategies are dependent on market value uncertainty. Specifically, there is a consensus that innovation speed (1) is most propitious in environments characterised by competitive intensity, technological and market dynamism, and low regulatory restrictiveness; (2) can be positively or negatively affected by strategic-orientation factors and organisational-capability factors; and (3) has an impact on development costs, product quality, and ultimately project success [9].

Meanwhile, recent developments in artificial intelligence have prompted a debate about the repercussions of innovations on labour markets. For example, one working paper [1] discussed the effects on labour markets in the United States. They highlighted the current concern about the increased replacement of workers by robots and computer technologies and the probable impact on the future of jobs and salaries [1]. They analysed the increase in the use of industrial robots during 1990 and 2007 on US local labour markets and created a model that represents the competition between robots and human labour in managing different activities. Their work demonstrated that robots are highly likely to reduce the available number of jobs and salaries, potentially pushing workers into unemployment. They estimated that "one more robot per thousand workers reduces the employment to population ratio by about 0.18–0.34 percentage points and wages by 0.25–0.5 percent" [1].

This is a critical concern, because the real or the potential value of a market derives from the quantity of jobs created and respective salaries. Individuals without jobs cannot buy products or services. This concern is being overlooked by executives.

Along a similar vein, another article [2] underlines the impact of artificial intelligence on the labour market. The big question is: "Will smarter machines cause mass unemployment?" This question speaks to the tendency to increase the levels and quality of automation and improve it, as well as the anxiety that workers now face.

The risk of replacement is not limited to unskilled workers but ranges across all sectors. Radiologists, doctors, or even lawyers, for example, now run the risk of seeing their expertise dispensed by computers. Unfortunately, this possible displacement is not on any political agendas. Instead, it is the owners of large companies who are sounding the alert of this systemic issue. The problem is not that all types of work in various sectors of activity are affected, but in routine work [2]. There is a simple explanation for it. In fact, there is a collusion of interests between companies and employees. Humans are not prepared to perform highly routine tasks with perfection. They become demotivated and the quality of products and services decreases over time. That is why robots can easily outdo humans in regard to quality standards, efficiency, and cost reduction. Machines can already perform many forms of routine manual labour and are now able to perform some routine cognitive tasks as well. This scenario is only like to increase. So we must ask: What are the jobs most susceptible to replacement by robots? A 2013 study of the likelihood of the computerisation of 702 occupations found that in the United States 47% of jobs were at risk of potential automation [10]. The researchers concluded that "recent developments in machine learning will put a substantial share of employment, across a wide range of occupations, at risk in the near future". Similar studies estimated at-risk jobs make up 35% of the workforce for Britain (where more people work in creative fields that are less susceptible to automation) and 49% for Japan [2]. These numbers give the dimension of "job polarisation" [11].

28.4 The INOV-I Model

The INOV-I Model is a multi-agent model built using the NetLogo software [12] intended to analyse the impact of innovation through artificial intelligence, where products or services are primarily produced by robots rather than people, and the impact of this innovation on market value. It is an adaptation from the work of [13].

In this particular model, companies represented by patches invest in robots in order to fulfil production orders or to provide services which they have in the pipeline. In these cases, workers can only find jobs in companies that are not completely automated. After multiple interactions, we can verify the impact of robotics on market value. It is also possible to identify the dilemma for executive managers: efficiency provided by robotics vs. loss in market value.

28.4.1 INOV-I Model Parameters

The INOV-I Model is composed of a set of parameters which characterises each type of company and their respective behaviour. Another set of parameters establishes the attributes of workers which influence their market behaviour.

Some parameters are controlled by the researcher and others were totally outside investigator control. In the second case, the parameters were settled randomly where their variations depended on dynamics arising from interactions between companies and agents or among agents.

This simple artificial economy is composed of two types of economic agents: companies and workers. Companies represented by NetLogo patches are labour suppliers, and workers are the agents who demand jobs in exchange for income. The symbol η represents the number of workers participating in this. As happens in real life, those agents have a set time period in which to carry out their tasks. The minimum age at which workers start in a company is given by the parameter "ma" and represents the youngest age at which a worker can legally enter into this market. On the other hand, the model sets the maximum legal retirement age at "MA". This age is the longest age (ticks) that an agent can work on this economy. This parameter has the goal to represent the restrictions imposed by the laws, in terms of retirement age.

Companies vary significantly in terms of innovation. Parameter ϕ indicates the initial proportion of companies which are completely robotic. Basically, these are the firms with best efficiency ratios in the market with high levels of quality in production and with zero defective products or high standards of services provided. Consequently, these are companies focus on reducing costs, even those arising from human resources.

Like [13, 14] independently of being workers or stockholders, both stakeholders receive an income represented by parameter φ. In truth, this parameter should be split off, since it represents distinct sources of money with different levels of taxation. For the purposes at hand, however, the data gathered under this parameter is sufficient considering the information needed. Income grows from time to time "due to dividend distributing, to attract better stakeholders, or even as a result of negotiations between government, labour unions and employers' associations, in a context of social agreement" [14]. The time interval of this growth is given by δ, and in accordance with [14, 15], this parameter determines the frequency of these negotiations. If the cycle is long, the labour market may be nearing full employment. If the cycle is overly short, it means the labour market is boiling up. Increases in income lead to a subsequent increase in market value. The parameter g represents growth during the interval of time and reveals the dependency of companies' productivity. Thus, there is an upper salary limit which each company can afford, and this, in turn, influences the value of this specific market. This maximum salary is represented by "ς" in this model. Parameter "sm" represents speed-to-market, while "td" indicates time to delivery. Another important attribute of workers is the fact that they are averse either to changing jobs, within companies, or to taking jobs at new companies which might require them to move far away, even if this move might enable them to find a job with more recognition of their value, relative to robots. Therefore, the model includes a parameter π that reflects the maximum mobility that workers are willing to accept. In truth and from another perspective, this parameter reveals the capacity of workers to adapt to new circumstances, when their work loses value vis-à-vis robots.

Parameter "c" represents workers' participation in the market as consumers, due to the income received. This parameter influences companies' market value. We must not forget the market is powered by workers' salaries.

Taxes are equally important in establishing market value. If income taxes over work are high, consumers' purchasing power is considerably reduced. If taxes are low, then consumption and consequently market value should rise. In modern societies, agents have the accountability to pay taxes, like what happens today in modern societies. The model provides for three levels of income taxes, low, average, and high, based on the Portuguese Tax System.

The INOV-1 Model accounts for workers' expectations of wealth, using, as did [15], a parameter called maximum wealth expectancy, ϖ. Workers may choose to move abroad when salary levels do not meet workers' expected consumption patterns and tax obligations.

28.4.2 INOV-I Model Description

The model was designed and built to analyse the impact of innovation arising from robotics and respective speed on market value. The model represents a simplification of a market economy and consists of companies that invest in robots to improve their levels of efficiency and to offer jobs. The number of workers depends on the innovative capacity of those companies that opted for this strategy. On the other hand, we have workers demanding jobs and investors who invest in potentially profitable companies from which they expect to receive dividends. Those agents have the goal to maximise their wealth during their working or investing life. For this reason, agents switch to companies offering better salaries, having higher value, or distributing attractive dividends, in order to increase their wealth. Therefore, at each turn, each agent consumes and saves part of his/her income, in order to improve his/her wealth (ω).

Companies are represented by NetLogo patches. A dark patch symbolises a company who opted to invest more in robotics and consequently does not need to hire employees. On the contrary, a light yellow patch symbolises a company lacking an innovative focus or without capacity to invest in robots. This company opts to use more workers, because of the scarcity of resources. However, the same company is normally less efficient than its competitors. It provides, however, greater potential for workers to find employment.

28.5 Preliminary Results

The results in this work came from experiments using version 5.0.4 of the NetLogo framework [12]. NetLogo is a programmable modelling environment that allows us to simulate natural and social phenomena. It is particularly recommended for modelling complexity and artificially reproduces these phenomena.

At this stage of research, the main goal is strictly committed with the scope of this article. This section presents the results of the simulations and then analyses their impact on different aspects of society. It next simulates the variation in speed of innovation and its impact on market value. Speed is measured during periods of innovation defined for this exercise to be 1 week, as time unit.

28.5.1 Initial Percentage of Innovative Companies

When we started to analyse the initial percentage of innovative companies that participate in this artificial economy, we can see that, when the percentage is small (about 5%), the proportion of innovative companies decreases and most workers in the market are fundamentally from low-income class. The percentage of market value increases from 0.214 in the first period of time to 0.455 in the latter simulations. It can also be seen that workers are more dispersed throughout the market (see Figs. 28.1, 28.2, 28.3, and 28.4).

However, when the percentage of innovative companies is high, we can identify greater numbers of medium-income workers and investors [34–345] in the market after 1240 months of simulation. It is also visible that workers are more concentrated in the market and that market value growth varies from 26.4% at the beginning of simulation to 39.3%.

It is evident that the impact on market value growth is higher, when the proportion of innovative companies is higher (25%). This suggests that the higher the investment in robotics, the lower the value of the market. Turbulence in the labour market is also obvious (see Figs. 28.1, 28.2, 28.3, and 28.4).

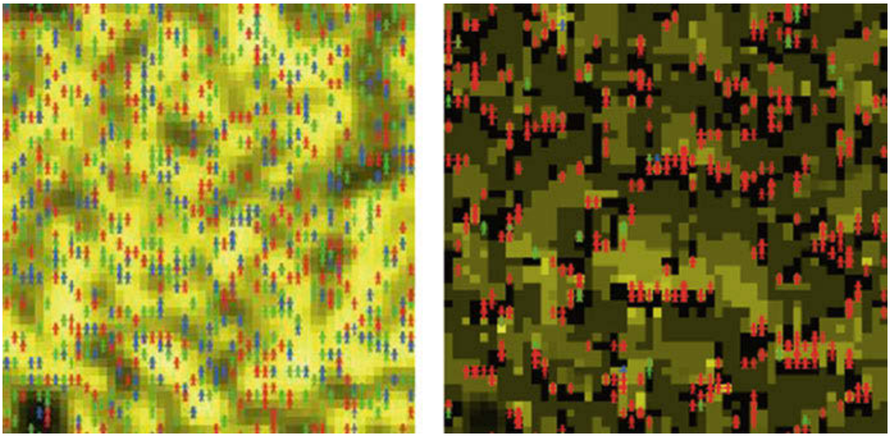


Fig. 28.1 Proportion of innovation – 5%

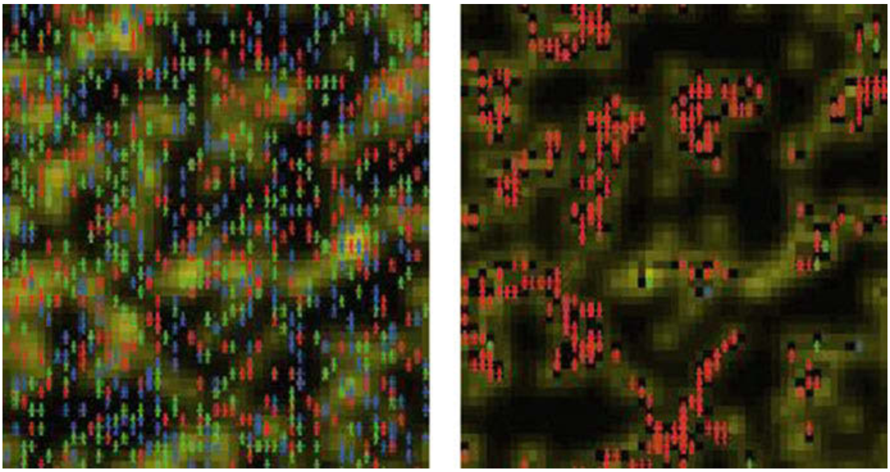


Fig. 28.2 Proportion of innovation – 25%

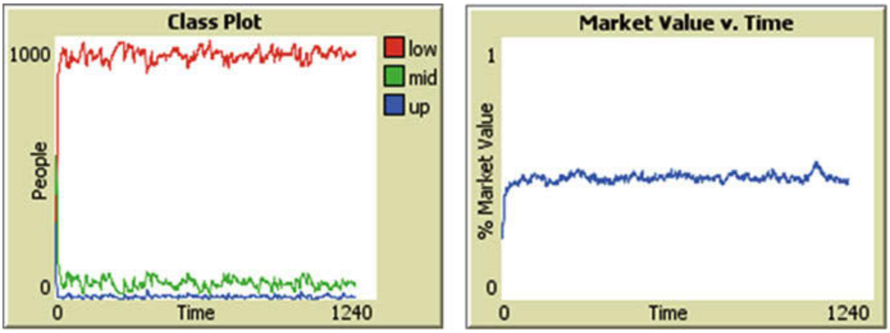


Fig. 28.3 Market value growth for 5% (From 21.4% to 45.5%)

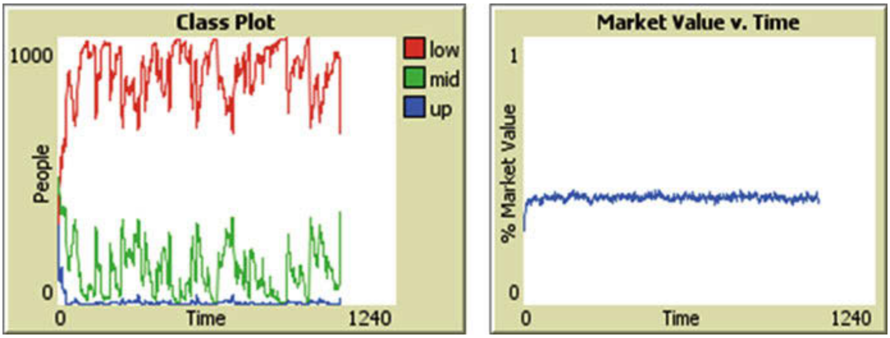


Fig. 28.4 Market value growth for 25% (From 26.4% to 39.3%)

28.5.2 Speed-to-Market Effect

As mentioned earlier, speed-to-market effect also influences the relation between robotic innovation and market value (Figs. 28.5, 28.6, 28.7, and 28.8).

Results for development intervals of 10 years demonstrated that (1) the lower the time-to-delivery innovation, the lower the number of jobs available and consequently the lower the market value and (2) most workers are from low income which reveals the value of the work under this scenario.

Another important output retrieved is the fact that market value grows more at the beginning of the simulation (short run) than in the long run (see Figs. 28.3 and 28.4 or 28.7 and 28.8) (Figs. 28.9, 28.10, 28.11, and 28.12).

Finally and regarding speed-to-market, we verified that the faster the implementation of innovation, the fewer the number of jobs created. This fact is perfectly clear in the simulations illustrated in Figs. 28.9 and 28.11. The artificial economy is seen to be completely black after simulation, and the growth of market value decreased over time. This result totally contradicts the previous ones. It is also important to call

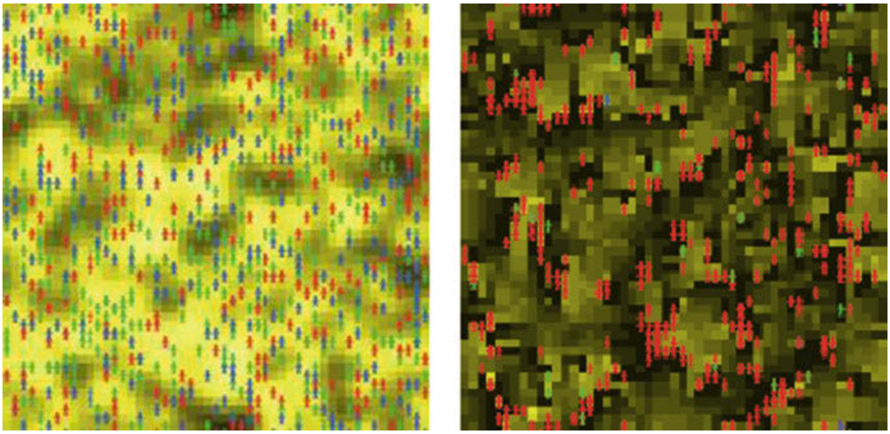


Fig. 28.5 Speed-to-market (Innovation developed interval 10 years; Time to delivery: 2 months; Innovation proportion: 12%)

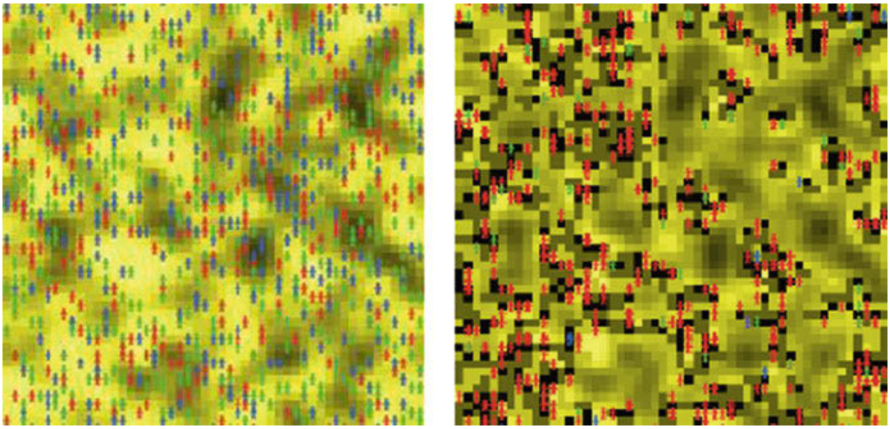


Fig. 28.6 Speed-to-market (Innovation developed interval 10 years; Time to delivery: 10 months; Innovation proportion: 12%)

attention to the rise in the number of workers/investors of medium and high income. These results prompt a hypothesis that the gains obtained from cost reductions and distributed by dividends may compensate for the loss in market value growth. It is a highly unique case.

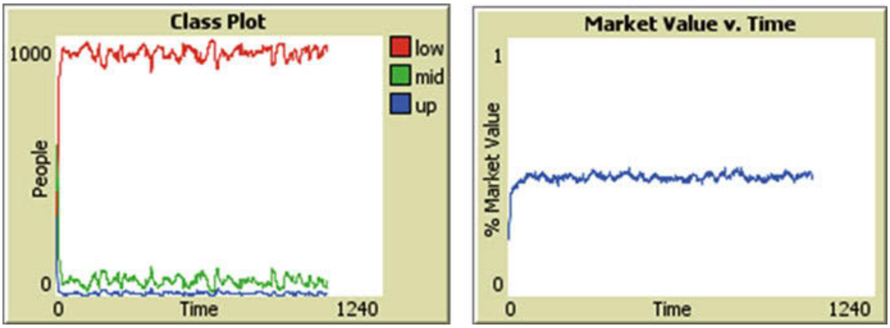


Fig. 28.7 Market value growth for TD 2 m (From 24.8% to 47.6%)

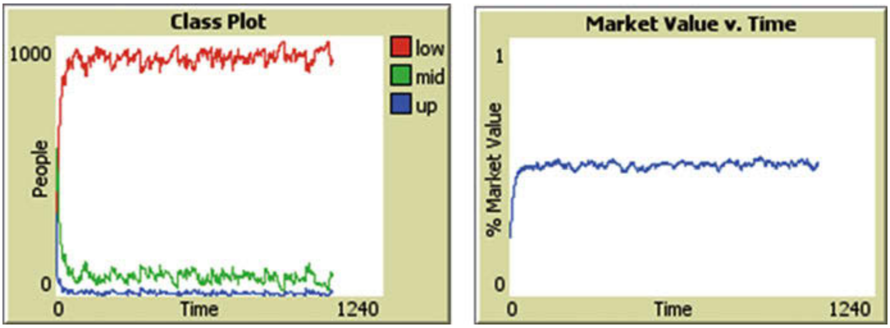


Fig. 28.8 Market value growth for TD 10 m (From 20.7% to 53.1%)

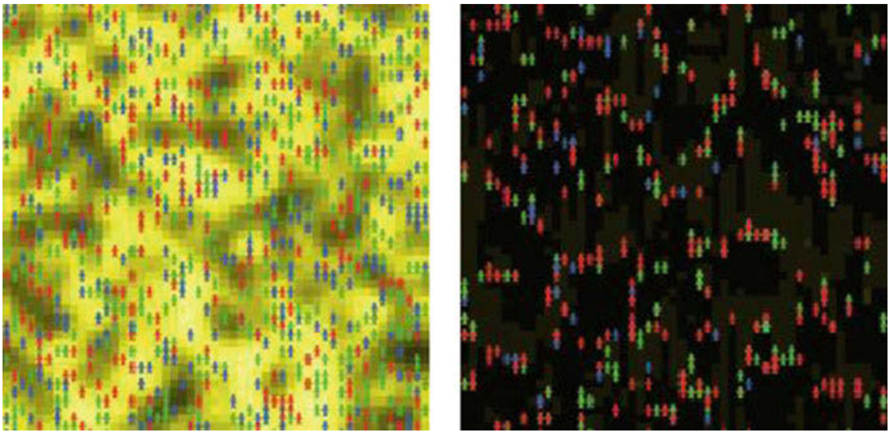


Fig. 28.9 Speed-to-market (Innovation developed interval 2 years; Time to delivery: 2 months; Innovation proportion: 12%)

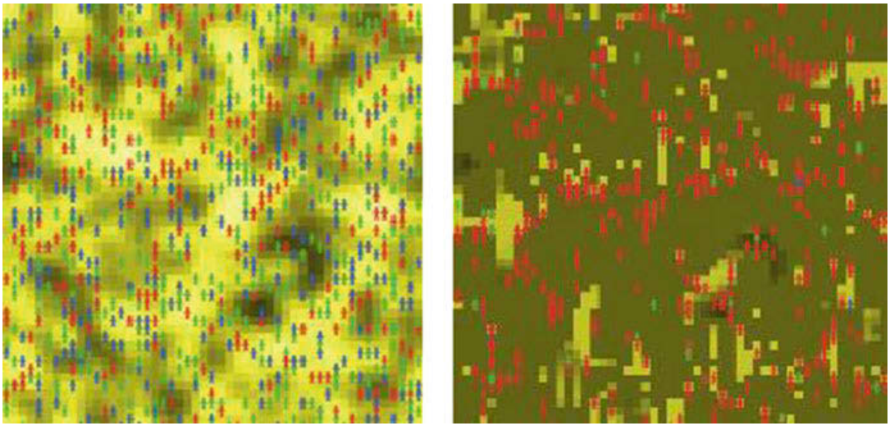


Fig. 28.10 Speed-to-market (Innovation developed interval 2 years; Time to delivery: 10 months; Innovation proportion: 12%)

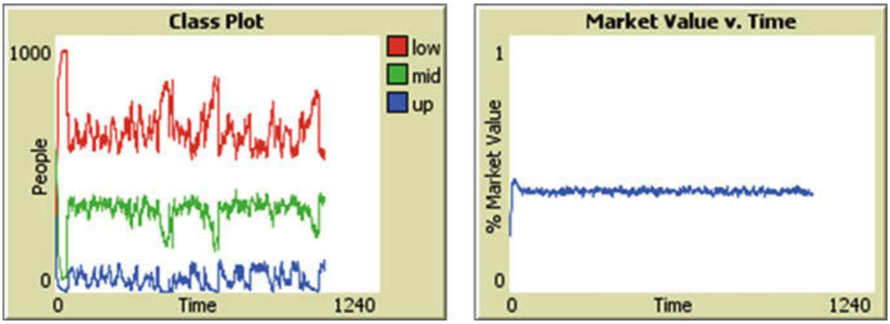


Fig. 28.11 Market value growth for TD 2 m (From 22.1% to 34.3%)

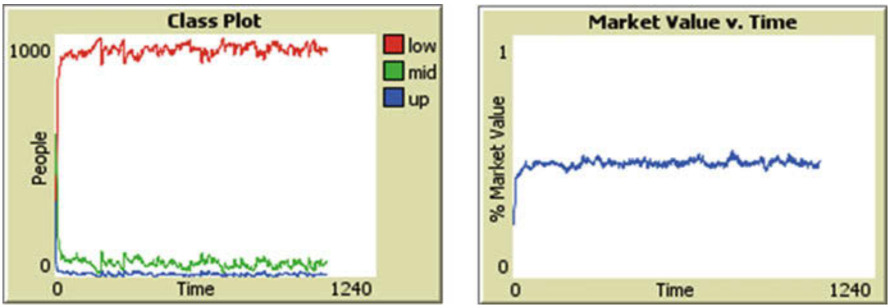


Fig. 28.12 Market value growth for TD 10 m (From 25.5% to 49.7%)

28.5.3 Mobility Effect

The next step was to simulate variations in the agent mobility parameter which indicates workers' willingness to relocate in search of a job far away from where they live. The obtained results revealed a positive correlation between workers' propensity to relocate and medium-income levels of workers/investors. The growth in market value is superior when mobility is high and inferior when the mobility is small (residual). A point to call attention to is that when mobility is limited, the number of companies investing in innovative robots is high (see Figs. 28.13, 28.14, 28.15, and 28.16).

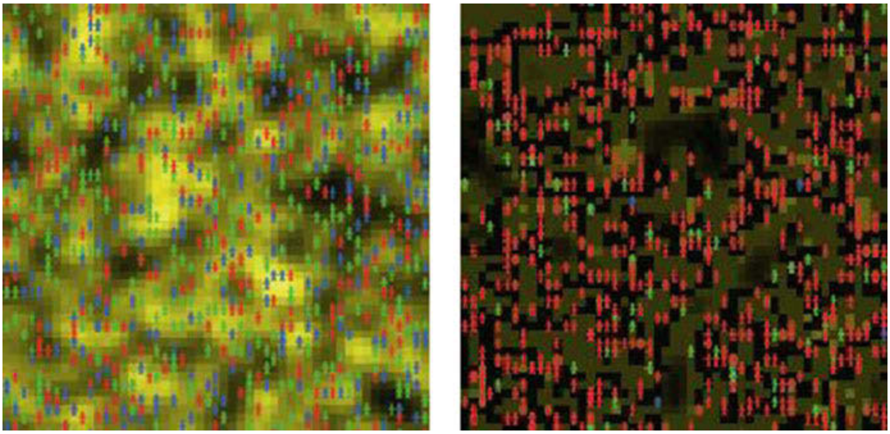


Fig. 28.13 Mobility effect (Mobility parameter: 1 patch)

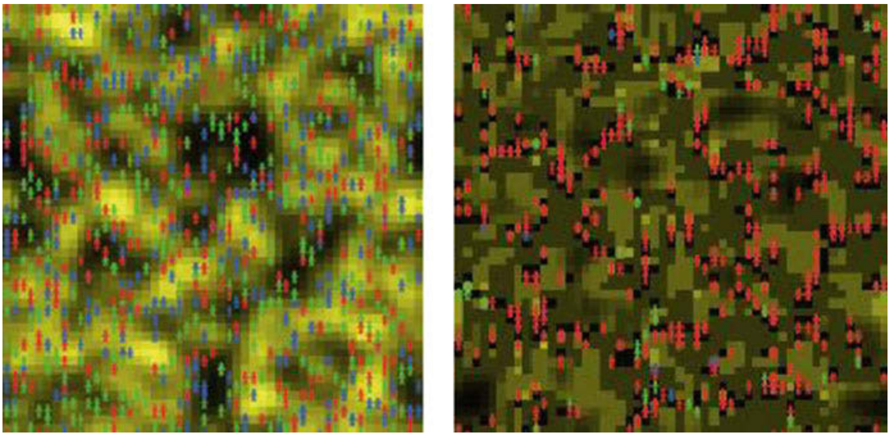


Fig. 28.14 Mobility effect (Mobility parameter: 15 patches)

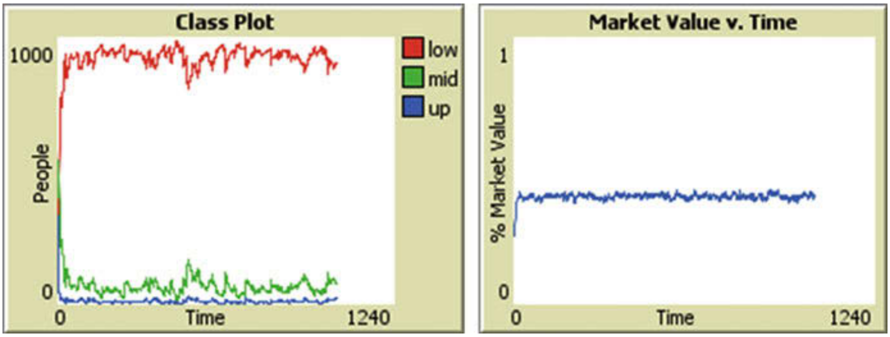


Fig. 28.15 Market value growth – 1 patch (From 27.6% to 40.7%)

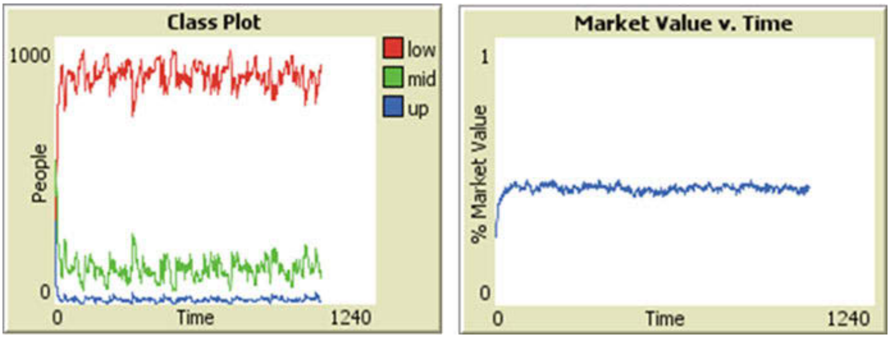


Fig. 28.16 Market value growth – 15 patches (From 24.6% to 42.1%)

28.6 Conclusion and Future Work

In this work, under the multi-agent-based systems methodology, a model was developed to study the indirect influence of innovation in robotics on the growth of market value for companies. Results have revealed a couple of interesting findings: (1) innovation through robotics negatively affects the market value of companies and there is no balance between innovation and companies' market value; (2) the higher the proportion of innovative companies, the lower the market value; (3) the relation between speed-to-market and market value is negative; and (4) the mobility effect interferes on a smaller scale with market value. Future research should address combined simulations of the various parameters, since the standard adopted was the ceteris paribus.

References

1. Acemoglu, D., Restrepo, P.: Robots and jobs: evidence from US labor markets, NBER working paper No. 23285 (2017)
2. Morgenstern, M.: Artificial intelligence: the impact on jobs, automation and anxiety. In The Economist, Special Report (2016)
3. Ford, M.: The Rise of the Robots. Technology and the Threat of Jobless Future. Basic Books, New York (2015)
4. Amit, R., Zott, C.: Value creation in e-business. Strateg. Manag. J. **22**(6/7), 493–520 (2001)
5. Amit, R., Zott, C.: Creating value through business model innovation. MIT Sloan Manag. Rev. **53**(3), 41–49 (2012)
6. MacCormack, A., Murray, F., Wagner, E.: Spurring innovation through competitions. MIT Sloan Manag. Rev. **55**(1), 25–32 (2013)
7. Ringel, M., Taylor, A., Zablit, H.: The most innovative companies 2015: four factors that differentiate leaders, The Boston Consulting Group Report (2015). https://media-publications.bcg.com/MIC/BCG-Most-Innovative-Companies-2015-Nov-2015.pdf
8. Chen, J., Reilly, R.R., Lynn, G.S.: The impacts of speed-to-market on new product success. IEEE Trans. Eng. Manag. **52**(2), 199–212 (2005)
9. Bhens, S., Lau, L., Markovitch, S.: Finding the speed to innovate, in Digital McKinsey, McKinsey & Company (2015). http://www.mckinsey.com/business-functions/digital-mckinsey/ourinsights/finding-the-speed-to-innovate
10. Frey, C., Osborne, M.: The future of employment: how susceptible are jobs to computerisation? Oxford Martin School, University of Oxford, Oxford OX1 1PT, United Kingdom Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, United Kingdom (2016). https://doi.org/10.1016/j.techfore.2016.08.019
11. Stiglitz, J.E.: The Price of Inequality: How Today's Divided Society Endangers Our Future, Kindle edn, p. 9. New York, W.W. Norton (2012)
12. Wilensky, U.: NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston (2012)
13. Magessi, N., Antunes, L.: Shadow economy and wealth distribution, in Artificial Economics 10th World Conference, Barcelona, post proceedings in Advances of Artificial Economics, Lecture Notes in Economics and Mathematical Systems, vol. 676, 2015, pp. 169–179. Springer (2014)
14. Magessi, N., Antunes, L.: Emigration or tax evasion? Solutions for Portuguese to escape from its economic crisis" in AAMAS 2015 – 14th International Conference of Autonomous Agents, Istanbul, 2015, proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems, pp. 1747–1748, International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC ©2015 (2015)
15. Magessi, N., Antunes, L: Comparing tax avoidance and tax evasion in terms of wealth distribution. In: 11th Artificial Economics Conference, in Advances of Artificial Economics, Springer (2015)

Chapter 29
Making Use of Fuzzy Cognitive Maps in Agent-Based Modeling

Sara Mehryar, Nina Schwarz, Richard Sliuzas, and Martin van Maarseveen

Abstract One of the main challenges in Agent-Based Modeling (ABM) is to model agents' preferences and behavioral rules such that the knowledge and decision-making processes of real-life stakeholders will be reflected. To tackle this challenge, we demonstrate the potential use of a participatory method, Fuzzy Cognitive Mapping (FCM), that aggregates agents' qualitative knowledge (i.e., knowledge co-production). In our proposed approach, the outcome of FCM would be a basis for designing agents' preferences and behavioral rules in ABM. We apply this method to a social-ecological system of a farming community facing water scarcity.

Keywords Knowledge co-production · Participatory modeling · Fuzzy cognitive mapping · Agent-based modeling

29.1 Introduction

Agent-Based Modeling (ABM) is a dynamic method for understanding and predicting the collective behavior of multi-agent systems, given the motives and preferences of individuals [1]. In principle, ABM requires specifying agents' available *actions*, behavioral *rules*, and decisions' *impacts* in each specific situation. To employ ABM in real-world applications, one main challenge is to formalize these three aspects (i.e., actions, rules, and impacts) such that the qualitative and

S. Mehryar (✉)
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, AE, The Netherlands

Grantham Research Institute on Climate Change and the Environment, London School of Economics (LSE), London, UK
e-mail: S.Mehryar@lse.ac.uk

N. Schwarz · R. Sliuzas · M. van Maarseveen
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, AE, The Netherlands
e-mail: N.Schwarz@utwente.nl; R.Sliuzas@utwente.nl; M.F.A.M.vanMaarseveen@utwente.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_29

quantitative knowledge and decision-making processes of stakeholders will be reflected. Many ABMs avoid addressing this challenge by relying on rational choice theory to describe their agents' behavior [2, 3]. However, stakeholders' behavior is usually not purely rational and often far more complex than assumed in such theories [4]. In particular, this is a problematic assumption in cases where preferences and decisions of agents highly depend on environmental dynamics, emerging social norms, and information accessibility. An alternative approach in such cases would be to inform ABM with participatory methods that collect qualitative data from stakeholders.

In this paper, we employ a Fuzzy Cognitive Mapping (FCM) method [5] to formulate and parametrize the qualitative knowledge gained by stakeholders (i.e., co-produced knowledge) and translate the FCMs to be used in ABM (Sect. 29.2). This methodology is demonstrated with the case of a farming community facing water scarcity.

29.2 Methodology

In general, FCM method enables collecting and representing stakeholders' knowledge about a particular problem [6]. More specifically, stakeholders' perception about influential variables and causal relations among these variables are represented in a directed graph structure. In this graph, variables appear as nodes (concepts) and causal relations as weighted links (connections) [7]. Below, we introduce a method that enables translating the knowledge represented in FCM for ABM development.

29.2.1 Collecting Knowledge

FCM models are usually developed with a participatory approach. Stakeholders who are familiar with the operation and behavior of the system or specific problem of the system are asked to mention the most important concepts (environmental, social, ecological, or economic variables), their causal relations, and weights of the connections (i.e., how much a change of one concept causes a change in another concept) [7]. To be able to use FCM results for ABM, stakeholders should also be asked about their (1) responses (actions) to the system or that specific problem and (2) causes and impacts of their actions from/on environment variables (conditions and impacts).

29.2.2 Representing Knowledge

The data gathered during the interviews can be categorized and represented in a graph structure as follows (Fig. 29.1):

1. Action concepts: which are the concepts mentioned by stakeholders as their responses to the system. For ABM use, they represent the set of possible actions that can be taken by agents. Size of concepts in FCM can be shown by the number of times they have been mentioned by stakeholders. Therefore, size of actions in FCM can represent *order* (*preferences*) of agents required in ABM.
2. Impact concepts: which are output concepts of each action along with their causal network, i.e., direct and indirect impacts of that action. Impact concepts are usually dynamic variables (with changing states), e.g., agriculture production, precipitation, or population change.
3. Condition concepts: which are input concepts of each action representing driving forces or causes of that action. Condition actions can be either dynamic—e.g., access to groundwater—or fixed (true/false) variables, e.g., having document or legal permission.
4. Driving connections: connections linking conditions and actions. These connections are not accompanied with causal weights, rather they represent implication and are interpreted as may "lead to" [8].
5. Impact connections: connections linking actions-impacts and impacts-impacts. These connections have causal weights, which reflect direct and indirect impact of actions on dynamic variables.

Having such information, for each action a set of Conditions-Action-Impacts (CAI) can be extracted from FCMs to be used in ABM development. In addition

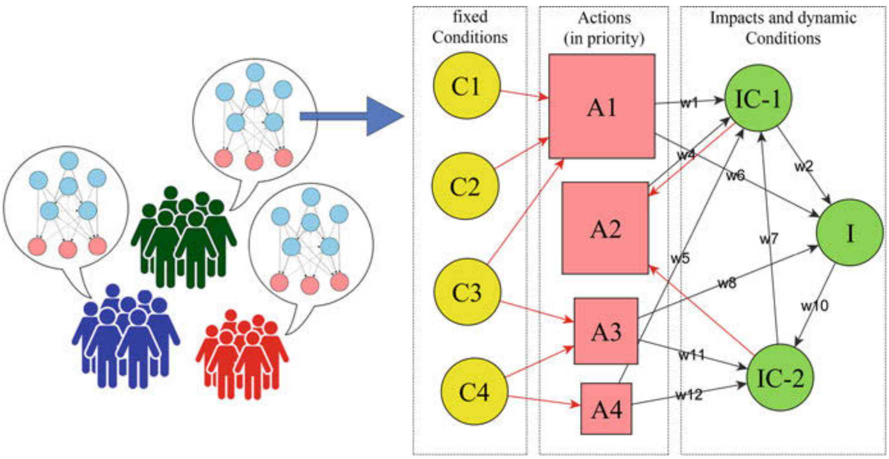


Fig. 29.1 Translating FCM model into the CAI map. Red and black arrows show driver and impact connections, respectively. *A* action, *C* condition, *I* impact

to the sequence of actions and their conditions and impacts, ABM development requires timing of certain actions (frequency of actions and one time actions vs repeating actions), randomness (in behavioral rules of agents), and spatial dimension (in case of varying spatial attributes). Since these aspects cannot be represented in a FCM, they should be added via quantitative data, complementary literature review, and local knowledge of experts collected during interviews.

29.3 Case Study and Preliminary Results

To illustrate the proposed methodology, we used the case study of a farming community facing water scarcity in Rafsanjan, Iran. Farmers take different kinds of adaptive actions (based on their social-spatial situation) to satisfy their water demand for pistachio production. Farmers' actions have different impacts (based on location and size of farms) on environmental properties as well as on other farmers decisions and actions. The main objective is to simulate the impact of aggregated farmers' adaptive actions on overall groundwater use in the region. For this objective, we used the FCM data collected in our previous study [9].

Therefore, the individual FCMs were developed by interviewing 60 farmers from different locations and social-economic situations. The farmers' knowledge about the main causes and impacts of water scarcity in their regions, their adaptive actions toward water scarcity, and influence of those actions on other variables of the system have been collected.

Although agents have the same preference, i.e., satisfying their water access, their decision-making mechanisms to achieve this goal are significantly different based on their economic situation. Therefore, the farmers' FCMs were developed within three groups of small, medium, and large farmers. The group-specific FCMs represent the set of farmers' actions to adapt with water scarcity in the order of farmers' preferences (node size in FCM). For example, large farmers' set of actions are *buying small farms* from medium and small holders, *desalination*, *purchasing water* from medium and small holders, *deepening wells*, *reducing farm's area*, and *relocating farms* in order. While medium farmers do not afford first three actions of large farmers, their set of actions include *deepening wells*, *integrated farming* with other medium farmers, *irrigation system change*, and *reducing farm's area* (Fig. 29.2). Small farmers have few options in their set of actions, which are basically *irrigation system change* or *turning off their well pumps* during the night or over winter. In addition to the set of actions and order of actions, CAI maps gave us the conditions for each action as well as the impact weights of each action on environment variables. As Fig. 29.2 shows, each action has condition concepts along with driver connections and a network of impact concepts along with impact connections. Thereby, the set of CAI for all actions have been extracted and combined with (1) time scale, (2) randomness, and (3) spatial diversity of the system to be used in an ABM model development. For example, randomness has been used for the actions with the same priority—e.g., *integrated farming* and *irrigation system change* in medium farmers' FCM—and time scale and

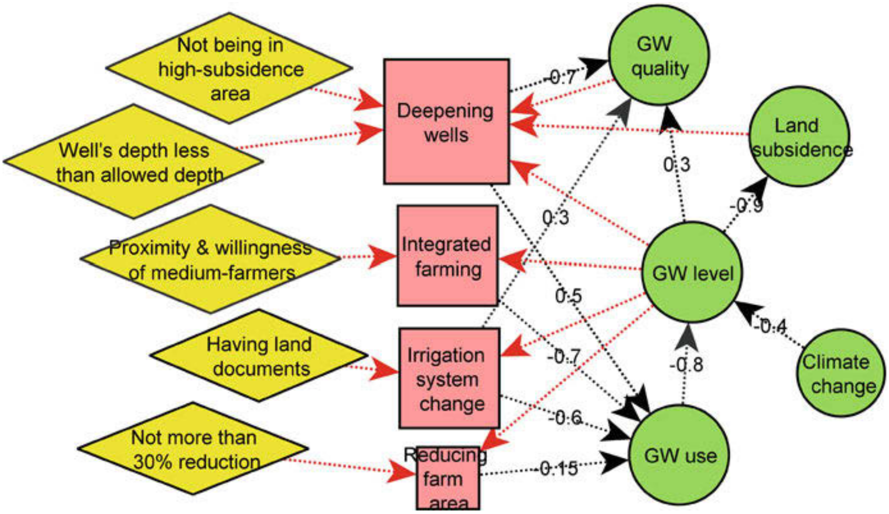


Fig. 29.2 CAI map of medium farmers based on their FCM. Red and black arrows show driver and impact connections, respectively. *GW* ground water

Table 29.1 CAI table of medium farmers

	Conditions	Actions	Impacts
Priority 1	Well depth < allowed well depth Farm location is not in high subsidence and poor GW quality areas	**Deepening wells**	Direct impacts: GW use Indirect impacts: GW level, land subsidence, and GW quality
Priority 2	Medium land in neighbor Neighbor medium-farmer is willing to integrate his/her farm Action has not been executed	**Integrating farming**	Direct impacts: GW use and quality Indirect impacts: GW level and land subsidence
	Farmer has land document Action has not been executed	**Irrigation system change**	Direct impacts: GW use and quality Indirect impacts: GW level and land subsidence
Priority 3	Land size >= 70% initial land size	**Reducing farm area**	Direct impacts: GW use Indirect impacts: GW level, land subsidence, and GW quality

spatial heterogeneity have been added to the conditions of actions to specify the frequency of actions, e.g., *integrating farm* as a one-time actions vs *deepening wells* that may happen several times before reaching the permitted wells depth, and the place of actions, e.g., *desalination* only happens in areas with poor quality of groundwater (Table 29.1). Preliminary results of this model show aggregation of groundwater use in different regions of this case study considering different farmers' actions, adaptations, and interactions with changing environment. Having the current situation of groundwater use by farmers, impacts of different policy alternatives can be simulated on changing overall groundwater use of region.[1]

29.4 Conclusion

We presented a method that enables translating qualitative co-produced knowledge (from FCM outputs) as an input for ABM development. Our proposed method includes aggregating individual interviews into group-specific FCMs and setting up CAI diagrams that provides inputs for ABM development. We also illustrated the applicability of this method using a case study in a farming community facing water scarcity. However, this method does not provide all information required in an ABM (e.g., temporal and spatial dynamics, stochasticity, ...). Therefore, quantitative and objective data (e.g., from literatures, reports, surveys, historical data) is complementary next to FCM to provide data for ABM.

Funding Sources This work was supported by Faculty of Geo-Information Science & Earth Observation, University of Twente, Netherlands, the Grantham Foundation for the Protection of the Environment, and the ESRC via the Centre for Climate Change Economics and Policy, United Kingdom (grant number: ES/R009708/1).

References

1. Gilbert, N.: Agent-based models. Sage Publication, USA (2008)
2. Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., Janssen, M.A., McAllister, R.R., Müller, B., Orach, K., et al.: A framework for mapping and comparing behavioural theories in models of social-ecological systems. Ecol. Econ. **131**, 21–35 (2017)
3. Groeneveld, J., Müller, B., Buchmann, C.M., Dressler, G., Guo, C., Hase, N., Hoffmann, F., John, F., Klassert, C., Lauf, T., et al.: Theoretical foundations of human decision-making in agent-based land use models–a review. Environ. Model. Softw **87**, 39–48 (2017)
4. Johnson, J., Ormerod, P., Rosewell, B., Nowak, A., Zhang, Y.C.: Non-equilibrium social science and policy. In: Non-equilibrium Social Science and Policy, pp. 1–17. Springer, Cham, Switzerland (2017)
5. Kosko, B.: Fuzzy cognitive maps. Int. J. Man-Mach. Stud. **24**(1), 65–75 (1986)
6. Papageorgiou, E., Kontogianni, A.: Using fuzzy cognitive mapping in environmental decision making and management: a methodological primer and an application. In: International Perspectives on Global Environmental Change. InTech, Rijeka, Croatia (2012)
7. Özesmi, U., Özesmi, S.L.: Ecological models based on people's knowledge: a multi-step fuzzy cognitive mapping approach. Ecol. Model. **176**(1–2), 43–64 (2004)

[1]Detailed description on implementation of this method and results of case study have been presented in [10, 11].

8. Elsawah, S., Guillaume, J.H., Filatova, T., Rook, J., Jakeman, A.J.: A methodology for eliciting, representing, and analysing stakeholder knowledge for decision making on complex socio-ecological systems: from cognitive maps to agent-based models. J. Environ. Manag. **151**, 500–516 (2015)
9. Mehryar, S., Sliuzas, R., Sharifi, A., Reckien, D., van Maarseveen, M.: A structured participatory method to support policy option analysis in a social-ecological system. J. Environ. Manag. **197**, 360–372 (2017)
10. Mehryar, S., Sliuzas, R., Schwarz, N., Sharifi, A., van Maarseveen, M.: From individual fuzzy cognitive maps to agent based models: modeling multi-factorial and multi-stakeholder decision-making for water scarcity. J. Environ. Manag. **250**, 109482 (2019)
11. Mehryar, S.: Participatory policy analysis in climate change adaptation: from individual perceptions to collective behaviour. University of Twente, Faculty of geo-Information Science and Earth Observation (ITC), Enschede, Netherlands (2019). https://doi.org/10.3990/1.9789036547260

Chapter 30
Policy Option Simulation in Socio-ecological Systems

Sara Mehryar, Richard Sliuzas, Nina Schwarz, and Martin van Maarseveen

Abstract Social-ecological systems (SESs) i.e., systems linking human behavior and environmental dynamics, are characterized as complex, dynamic, and heterogeneous systems. A comprehensive policy simulation in SESs has to consider all these properties. In this paper, we present a conceptual analysis on key aspects of policy simulation in SESs and introduce a combined use of fuzzy cognitive mapping (FCM) and agent-based modeling (ABM) methods as an approach to cover those aspects. We illustrate the applicability of this combined method for policy simulation in the case of a farming community facing water scarcity.

Keywords Social-ecological systems · Policy making · Fuzzy cognitive mapping · Agent-based modeling

30.1 Introduction

In a broad perspective, social-ecological systems (SESs) are characterized as complex adaptive systems [1] in which wicked problems [2] may arise. While *complexity* refers to being nonlinear, causally interactive, heterogeneous, and temporally dynamic [3], *wickedness* refers to data-scarce/data-uncertain, multivariable, and multi-stakeholder issues [4, 5]. Accordingly, policy making in SES has to consider

S. Mehryar (✉)
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, AE, The Netherlands

Grantham Research Institute on Climate Change and the Environment, London School of Economics (LSE), London, UK
e-mail: S.Mehryar@lse.ac.uk

R. Sliuzas · N. Schwarz · M. van Maarseveen
Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, AE, The Netherlands
e-mail: N.Schwarz@utwente.nl; R.Sliuzas@utwente.nl; M.F.A.M.vanMaarseveen@utwente.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_30

all these features for a holistic policy option analysis prior to their implementation in the real world [3].

Therefore, based on our experience in SESs and following complex systems' and wicked problems' literature [1–3, 6], we categorized three main aspects that are essential to be considered in an effective policy analysis in SESs: (1) including consensus knowledge of stakeholders and their acceptance in policy simulation (2) considering individual and spatial heterogeneity of SESs, and (3) taking into account the time scale, causal relationships, and feedback loops of social-ecological interactions.

In this paper, we first present these three main aspects. Second, we introduce a combined use of fuzzy cognitive map (FCM) and agent-based modeling (ABM) as an approach of policy option simulation that covers all these aspects.[1] Finally, relying on the results of a 4-year project, we present findings of policy option analysis in a SES with combination of these two methods.

30.2 Conceptual Analysis

30.2.1 Consensus Knowledge

In general, using consensus knowledge of stakeholders in policy simulation—known as knowledge co-production—is a collaborative process that integrates (1) different stakeholders' knowledge and perceptions of their environmental issues and (2) possible stakeholders' response (e.g., acceptance and rejection) to that policy [5, 9]. This is particularly useful in policy simulation for SESs' *wicked problems*, i.e., ill-formulated problems because of their multivariables/multi-stakeholders environment and lack of available data for the whole system [4, 5]. Moreover, knowledge co-production helps to cover divergent knowledge of stakeholders with strong conflicting interests, i.e., another feature of SESs' wicked problems [5, 9].

Hence, using knowledge co-production in policy simulation may result in improving policy effectiveness by considering important features of complex and wicked SES problems, i.e., lack of data, multivariable environment, and multi-stakeholders with conflicting interests.

[1] See [7, 8] for a detailed discussion on technical aspects of our proposed approach. A comprehensive framework that integrates the conceptual and methodological aspects will be presented in the long version.

30.2.2 Individual and Spatial Heterogeneity

Individual heterogeneity refers to various types of involved stakeholders in SESs and highlights their different preferences, available actions, and long-term goals [10]. Moreover, spatial heterogeneity refers to the various environmental properties in different locations [1]. In policy making for SESs, it is important to not only aggregate knowledge of heterogeneous stakeholders but also represent their heterogeneity in policy impact analysis. Since, impacts of different policies may vary in different locations and on different individuals. For example, in a water scarce situation, different groups of farmers may have different adaptive actions based on their locations and social-economic situation, i.e., from buying water and irrigation system change (expensive options) to shrinking farm area (affordable option) or well deepening where aquifer situation allows (location-specific option). Considering social-spatial varieties in such cases results in a policy toolbox that corresponds to the individual and spatial heterogeneity of SESs.

30.2.3 Time-Scale, Causal Relationship, and Feedback Loop

Next to the heterogeneity, these are the main features of complex systems—including SESs:

- In a SES, there are different time scales in variables' changes and agents' actions. For example, slowly changing variables (e.g., population change) vs fast-changing variables (e.g., government policies) or high-frequency actions (e.g., farm irrigation) and low-frequency actions (e.g., buying lands).
- Causal relationships in human-environment interactions represent how a network of social and ecological factors influences humans' decisions and actions and vice versa. For instance, if farmers start to adapt a new irrigation technology, how much this adaptive action impacts on their groundwater use at first place and thereby, on overall groundwater level and production of region, and eventually on vulnerability and emigration of farmers. Capturing causality of SES variables is not possible via empirical correlations or a snapshot of causal relationships at one moment in time [12].
- Feedback loops in SESs are fundamental to understand both direct interaction of agents, i.e., how individual-level behavior and actions influence other individuals' actions, and indirect interaction of agents, i.e., how the outcome of performed behavior or actions impacts other factors of SESs which indirectly reinforce the same or other individual actions.

Considering these three properties in a policy simulation results in capturing the dynamicity of complex SESs.

30.3 Policy Simulation Models and Results

In this section, we illustrate how the consensus knowledge of stakeholders can be captured using a participatory method, i.e., fuzzy cognitive map (FCM). Then we show how social-spatial heterogeneity can be addressed using a dynamic modeling method, i.e., agent-based modeling (ABM). Finally, we present how a combined use of FCM and ABM methods can capture all important aspects of policy simulation in SESs. To illustrate the applicability of these methods, we rely on the results of a 4-year project that studies a farming community facing water scarcity in Iran [11].

30.3.1 Fuzzy Cognitive Mapping and Aggregated Knowledge

FCM uncovers the causal relationships between social and ecological factors by representing the SES with nodes (concepts), weighted links (connections), and many feedback loops. Methodologically, it collects the rich qualitative knowledge of stakeholders, aggregates them, and represents a semiquantitative outcome. This makes FCM an appropriate methodology for knowledge co-production. We developed 2 FCM models of 40 policy makers and 60 farmers in a farming community, to capture their knowledge of causes and impacts of water scarcity as well as farmers' adaptive actions toward water scarcity [4]. Then, the impacts of different policy options (e.g., economic diversification, technology adaptation, and monitoring policies) on the SES were simulated. Results of FCM model and simulation in this case study showed that among four policy options suggested by the government in Rafsanjan, farmers strongly believe in the impact of economic diversification on reducing water shortage, whereas the policy makers focus on the role of government control and monitoring policies to deal with water scarcity [4]. This methodology was useful for knowledge co-production in data-scarce situations as well as comparing the acceptability of different policy options. However, it does not represent the individual heterogeneity nor the spatiotemporal dynamics of the SES.

30.3.2 Agent-Based Modeling and Interactive Agents

Farmers' decisions and adaptive actions toward water scarcity in Rafsanjan can be very different based on their location and economic situation (heterogeneity). For example, large farmers tend to expand their lands and production by buying small farmers' lands and water. On the other hand, medium farmers look for collaborative farming and irrigation mechanisms with other medium farmers to increase efficiency of their groundwater use. Finally, small farmers prefer to sell off or shrink (dry off) their lands to adapt with water shortage. In this case, ABM [1, 13, 14] has been used to model farmers' collective behavior and simulate its impact on overall groundwater use. However, ABMs often have the problem of identification and justification of agents' behavioral rules as well as the causal relationships and nonlinear feed backs in human-environment interactions [3, 12]. Hence, we used the FCM's output knowledge for ABM model justifications and to provide causal relationships and feedback loops of SES for ABM development.

30.3.3 Combined FCM and ABM Methods and Results

While FCM covered consensus knowledge and ABM enabled modeling our heterogeneous farming community, only a combination could cover the last aspect, i.e., time scale, causal relationship, and feedback loop. Time scale in agents' actions and environment variables could be represented by ABM, whereas causal relationships and feedback loops in social-ecological interactions were shown by FCM. Moreover, with this combined method, we could simulate impact of policies considering preferences, interactions, and dynamic response of farmers over time which resulted in different outcomes comparing with those of FCM simulation. For example, real impact of the irrigation system change policy, i.e., subsidizing new irrigation system for farmers, depends on priority of this action for farmers comparing to their other possible adaptive actions. Results of policy option simulations with a combined FCM-ABM modeling methodology showed that (1) policy options (i.e., government control and monitoring, adapting new technologies, and participatory water management) have different long-term impacts in different locations of our case study and, (2) overall, policy of facilitating people participation in management and control of their groundwater use has the highest impact in reducing groundwater use. Surprisingly, adapting new irrigation technologies does not have any significant impact on reducing overall groundwater use in the region.

30.4 Conclusions

Policy option simulation in SES has to consider various features of complex adaptive systems and wicked problems. It is essential for an effective policy analysis in a SES to (1) include consensus knowledge of stakeholders and their acceptance in policy simulation, (2) consider individual and spatial heterogeneity of SESs, and (3) take into account the time scale, causal relationships, and feedback loops of social-ecological interactions. FCM can cover the first aspect, ABM the second, and only combination of these two can provide all three aspects. However, some social-political aspects, e.g., cultural differences and power relations, can be investigated in future studies.

Funding Sources This work was supported by Faculty of Geo-Information Science & Earth Observation, University of Twente, Netherlands, the Grantham Foundation for the Protection of the Environment, and the ESRC via the Centre for Climate Change Economics and Policy, United Kingdom (grant number: ES/R009708/1).

References

1. Filatova, T., Verburg, P.H., Parker, D.C., Stannard, C.A.: Spatial agent-based models for socio-ecological systems: challenges and prospects. Environ. Model. Softw. **45**, 1–7 (2013)
2. Xiang, W.N.: Working with wicked problems in socio-ecological systems: awareness, acceptance, and adaptation. Landsc. Urban Plan. **110**, 1–4 (2013)
3. Levin, S., Xepapadeas, T., Crépin, A.S., Norberg, J., De Zeeuw, A., Folke, C., Hughes, T., Arrow, K., Barrett, S., Daily, G., et al.: Social-ecological systems as complex adaptive systems: modeling and policy implications. Environ. Dev. Econ. **18**(2), 111–132 (2013)
4. Mehryar, S., Sliuzas, R., Sharifi, A., Reckien, D., van Maarseveen, M.: A structured participatory method to support policy option analysis in a social-ecological system. J. Environ. Manag. **197**, 360–372 (2017)
5. Olazabal, M., Chiabai, A., Foudi, S., Neumann, M.B.: Emergence of new knowledge for climate change adaptation. Environ. Sci. Policy **83**, 46–53 (2018)
6. Liu, J., Dietz, T., Carpenter, S.R., Alberti, M., Folke, C., Moran, E., Pell, A.N., Deadman, P., Kratz, T., Lubchenco, J., et al.: Complexity of coupled human and natural systems. Science **317**(5844), 1513–1516 (2007)
7. Mehryar, S., Schwarz, N., Sliuzas, R., van Maarseveen, M.: Making use of fuzzy cognitive maps in agent-based modeling. In: Verhagen, H., et al. (eds.) Advances in Social Simulation, Springer Proceedings in Complexity. Springer Nature, Switzerland (2020)
8. Mehryar, S., Sliuzas, R., Schwarz, N., Sharifi, A., van Maarseveen, M.: From individual fuzzy cognitive maps to agent based models: modeling multi-factorial and multi-stakeholder decision-making for water scarcity. J. Environ. Manag. **250**, 109482 (2019)
9. Gaddis, E.J.B., Falk, H.H., Ginger, C., Voinov, A.: Effectiveness of a participatory modeling effort to identify and advance community water resource goals in st. albans, vermont. Environ. Model. Softw. **25**(11), 1428–1438 (2010)
10. Macy, M.W., Willer, R.: From factors to factors: computational sociology and agent-based modeling. Ann. Rev. Sociol. **28**(1), 143–166 (2002)
11. Mehryar, S.: Participatory policy analysis in climate change adaptation: from individual perceptions to collective behaviour. University of Twente, Faculty of geo-Information Science and Earth Observation (ITC), Enschede (2019). https://doi.org/10.3990/1.9789036547260
12. Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., Janssen, M.A., McAllister, R.R., Müller, B., Orach, K., et al.: A framework for mapping and comparing behavioural theories in models of social-ecological systems. Ecol. Econ. **131**, 21–35 (2017)
13. An, L.: Modeling human decisions in coupled human and natural systems: review of agent-based models. Ecol. Model. **229**, 25–36 (2012)
14. Matthews, R.B., Gilbert, N.G., Roach, A., Polhill, J.G., Gotts, N.M.: Agent-based land-use models: a review of applications. Landsc. Ecol. **22**(10), 1447–1459 (2007)

Chapter 31
An Integrated Model to Assess the Impacts of Dams in Transboundary River Basins

Kavin Narasimhan, Nigel Gilbert, and Corinna Elsenbroich

Abstract This extended abstract presents an integrated agent-based and hydrological model to explore the impacts of dams in transboundary river basins where riparian nations have competing water uses. The purpose of the model is to explore the effects of interactions between stakeholders from multiple levels and sectors on the management of dams and its subsequent effects on the water-energy-food-environment (WEFE) nexus in river basins.

Keywords Dams · Coupled human-natural systems · Transboundary water management · Agent-based modelling · Hydrological modelling

31.1 Introduction

Dams are a critical component of water resources management in river basins. There are over 800,000 dams across the world, and 3700 dams are currently being planned in developing nations [1, 2]. These engineered systems often serve multiple purposes such as energy production, irrigation, water supply for industrial and domestic uses, downstream navigation, recreation, etc. In doing so, dams embed infrastructure into the water, energy, food, and environment nexus. On the other hand, by storing, diverting, and controlling the flow of water, which in turn affects sediment and nutrient transport in rivers and alters water temperature and chemistry, dams pose limitations that could negatively impact natural ecosystems, biodiversity, and the people whose livelihood depend on them [3]. Growing concerns about climate change and natural resources depletion have motivated decision-makers and authorities to prioritize the environmental and social impacts of dams.

The beneficial and negative impacts of a dam are governed by the interactions between dam structure (engineering dimension), its operation and management

K. Narasimhan (✉) · N. Gilbert · C. Elsenbroich
Centre for Research in Social Simulation, University of Surrey, Guildford, UK
e-mail: k.narasimhan@surrey.ac.uk; n.gilbert@surrey.ac.uk; c.elsenbroich@surrey.ac.uk

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_31

driven by multisectoral socioeconomic needs and geopolitical factors (human dimension), and the hydrology (interaction of water with the physical and biological environment through processes such as precipitation, evaporation, infiltration, etc.) of watercourses (natural dimension). The impacts of dams vary across sites based on the nature and characteristics of the engineered, human, and natural systems unique to each site [3]. Understanding these relationships, which impose profound, complex, and multiple societal and ecosystems impacts, is essential for stakeholders and decision-makers to plan the management of water resources in an integrated and sustainable way. Computational modelling offers one way to plan the management of environmental resources, especially in the context of being useful for informing policy making [4].

31.2 Modelling Water Resources Management: A Short Review

The common approach of modelling water resources as scarce goods and ascribing economic value to them has been widely used to inform and assess water management policies, e.g., hydro-economic modelling [5]. Although monetizing in this way is often useful to convert a multi-objective management problem into a single-objective problem (e.g., profit maximization), it does not allow assessing the qualitative or hard-to-quantify aspects of water demands, management methods, and institutions, e.g., issues related to environmental and ecological values and benefits, social equity, and the interdependencies between ecological, hydrological, and social systems [3, 5, 6]. A purely economic assessment also does not consider the inherent complexity of integrated human-environment systems such as dams.

On the other hand, agent-based modelling (ABM) is a method useful for representing and simulating the inherent complexity of coupled socio-ecological systems, which have interdependent natural and social components. It does so by defining and implementing complexity mechanisms such as nonlinearity, feedback, uncertainty, causality, and heterogeneity [4, 7]. These capabilities have contributed to the continued popularity of ABM as a viable approach to investigate the effect of decisions and strategies for managing coupled socio-ecological systems (also referred to as coupled human-natural systems or coupled human-environment systems) or more specifically socio-hydrological systems (coupled human-water systems) [7–9].

There are two main ways in which ABM has been used in the context of modelling human-water interactions. Firstly, agent-based models are intended as a form of *science-policy/modeller-expert interface tools* that modellers (e.g., academics), decision-makers (e.g., policy makers), stakeholders (e.g., actors in the river basin), system experts, and domain experts can develop and validate collaboratively, making sure everyone develops a shared understanding of the system and the issues being examined [10, 11]. Secondly, in an *integrated model* that includes other modules (e.g., hydrological, economic, engineering, etc.) and techniques (e.g., optimization, Bayesian belief networks, etc.), agent-based models are used to (1) model bottom-up decision-making, learning, and adaptation of actors, (2) consider nonlinear relationships between system components, (3) model spatially explicit representations of systems and their evolution over time, (4) represent self-organization of systems, and (5) model emergence of macro-level phenomena from micro-level behavior. It is also possible to combine the two applications, whereby an agent-based model is developed in a participatory context involving the relevant stakeholders, and is integrated with other social science (e.g., economic) and/or physical science (e.g., hydrological) modules to emulate specific real-world applications [6, 12]. ABM has also been used to generate scientific evidence that can be validated by stakeholders and experts using participatory processes [13].

In the following section, we present an integrated model to explore the effects of stakeholder interactions on the management of dams and its subsequent effects on the water-energy-food-environment (WEFE) nexus in the Volta River Basin (VRB) in West Africa. The novel aspect is the focus on the human dimension, which complicates the management of dams owing to the many layers of social, political, and economic institutions involved in river basin management and manifold end users even at the sub-basin level, e.g., small-scale farmers, large-scale farmers, hydro-electric power companies, industrial users, tourism and recreation users, and municipal water users [14–16]. As conflicts in river basins mostly arise at regional or local levels, which could then escalate into national conflicts, it is important to involve stakeholders from multiple levels and sectors when formulating water management decisions [17]. Decision-making should result from the lateral interactions between riparian nations and the vertical interactions starting from grassroots stakeholders (e.g., farmers, herders, village chief, etc.) through to national agencies (e.g., water resource commissions) and international agencies (e.g., river basin authorities) [18]. Another novel aspect of the model is the implementation described in the following section.

31.3 An Integrated Model to Explore the Impacts of Dams

The resources of the VRB, which has a transboundary watershed of approximately 400,000 km^2, are shared by six riparian countries: Côte d'Ivoire, Ghana, Togo, Benin, Burkina Faso, and Mali. A major part of the watershed lies in the coastal Ghana and land-locked Burkina Faso (around 42% each), while a minor part lies in the remaining four countries [19]. The differing economic priorities of the riparian countries in the VRB have an important role in the water management of the basin, e.g., Burkina Faso relies heavily on irrigation for farming, whereas Ghana's increasing electricity demand has made hydropower generation a priority. At the basin level, domestic water supply and irrigation for farming are key priorities for water [20]. The Volta Basin Authority (VBA) water charter was signed by the VRB riparian nations in 2006 as a means to support collaborative water management, enable coordination to resolve transboundary water conflicts and tensions, influence joint development, and achieve policy harmonization. However, the VBA is yet to become fully functional in terms of achieving its mandates and objectives [18, 21].

The differing economic priorities of the riparian nations and the lack of adequate governance mechanisms are both critical issues affecting transboundary water management in the VRB. Therefore, we conceptualize an integrated model to explore the possibilities of coordination among stakeholders in making water management decisions by considering their competing sectoral priorities and preferences. We consider the following sectoral priorities relevant in the context of the VRB: irrigation, hydropower generation, domestic water supply, and ecosystem health. In terms of preferences, we consider stakeholders' willingness to cooperate with the requests of other stakeholders in the system and contribute to capacity building in the basin, which are both important to avoid water conflicts [22]. The element of capacity building we consider is knowledge exchange between stakeholders which has been shown to improve the confidence of stakeholders, overcome mistrust, and enable cooperative action [17].

The integrated model illustrated in Fig. 31.1 has a hydrological model and an agent-based model. The former simulates the impact of stakeholders' dam operation decisions on the quality and quantity of water in the river basin. The latter consists of two layers: a network layer and a decision layer. The network layer defines the relationship between a diverse set of local, regional, national, and international stakeholder agents. The decision layer simulates the decision-making of individual stakeholders using a rule-based system. Below is a high-level description of each of these components.

The agent-based modelling component controls stakeholders' decisions associated with dam operations on a monthly time step which is a reasonable timescale for water management decisions [13]. First, a prespecified network configuration is

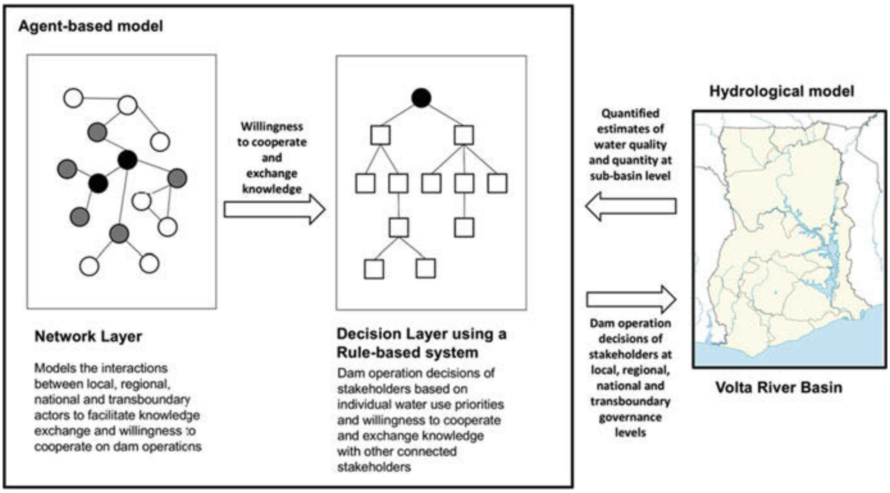


Fig. 31.1 An integrated model to assess the impact of the management of dams in the VRB

used to model the interactions between a diverse group of stakeholders with regard to willingness to cooperate on dam operations and knowledge exchange on water management issues. The VRB is divided into geographically and politically similar sub-regions, each of which is influenced or governed by a local stakeholder agent. Local stakeholders who have common water use needs are grouped under a regional stakeholder agent. The main water allocation and management bodies in each of the six riparian nations are the national stakeholders. The VBA and other relevant actors (e.g., financial institutions and NGOs) are the international stakeholders [18, 21].

Based on the interactions with other connected stakeholders in the system, each stakeholder agent makes a dam operation decision (or a decision that influences a dam operation decision of a higher-level stakeholder in its network). Decisions are derived using a rule-based system, developed by organizing real-world stakeholder inputs within the modelling framework, which in turn improves the effectiveness of the participatory approach used to develop and refine the model. A rule-based system includes (1) a working memory initialized with a set of facts relevant to the beginning state of each agent, (2) a rule set describing the actions to be performed under different conditions, (3) a matching scheme to decide which rules are applicable, and (4) a conflict resolution scheme to choose the most applicable condition-action rule among alternatives. The rule-based system enables each stakeholder agent to make appropriate dam operation decisions for different sectoral priorities (irrigation, hydropower generation, domestic water supply, and ecosystem health) based on individual targets (set based on historical data) and interactions with other stakeholders in the system.

The dam operation decisions of stakeholders are subsequently fed into a hydrological model developed using the Soil and Water Assessment Tool (SWAT[1]), free software used to explore and quantify the effects of human intervention (in our case, dam operation decisions) on the quality and quantity of water resources in a river basin. The outputs of the hydrological model (discharge, nutrient data, sediment data, crop yields at the sub-basin level, etc.) are fed into the agent-based model, which in turn affects the priorities and thereby the dam operation decisions of stakeholders at the next time step. In the last time step of a model run, a summary is generated describing the performance of the different sectoral priorities and instances of successful cooperation and knowledge exchange among stakeholders at different levels of governance. The model can be applied to different scenarios of stakeholder interaction (with regard to cooperation and knowledge exchange) and decision-making (with regard to dam operations based on individual water use needs and interactions with connected stakeholders) to assess the impact of dam operations on the WEFE nexus in VRB.

The inputs required for the integrated model are as follows. (1) The configuration of the VRB stakeholder network and data about the nature of the interactions between stakeholders are obtained from existing literature in the first instance. (2) The rule sets and priorities for the rule-based system specified are based on available

[1] https://swat.tamu.edu/

data (e.g., the Transboundary Waters Assessment Programme database,[2] the Global Reservoir and Dam database,[3] the AQUASTAT database[4]) and stakeholder inputs. (3) The SWAT hydrological model requires input data on weather, soil properties, topography, vegetation, etc., most of which are commonly available as open source datasets (e.g., the WaterBase dataset[5]).

31.4 Discussion

This paper presented an integrated model to be useful for exploring the impacts of multi-level stakeholder interactions and decision-making related to dam operations on the WEFE nexus in the Volta River Basin. We will implement the integrated model by linking a hydrological model developed using SWAT and an agent-based model developed using NetLogo.[6] We will use the JESS rule engine[7] to implement the rule-based system governing the decisions of stakeholders. Although the proposed implementation is specific to the VRB, the integrated model can be easily adapted to other river basins.

References

1. Kennedy, T.A., Muehlbauer, J.D., Yackulic, C.B., et al.: Flow management for hydropower extirpates aquatic insects, undermining river food webs. Bioscience. **66**, 561–575 (2016)
2. Poff, N.L., Schmidt, J.C.: How dams can go with the flow. Science. **353**, 1099–1100 (2016)
3. Bergkamp, G., McCartney, M., Dugan, P., et al.: Dams, ecosystem functions and environmental restoration. Thematic Rev. II. **1**, 1–187 (2000)
4. Zellner, M.L.: Embracing complexity and uncertainty: the potential of agent-based modeling for environmental planning and policy. Plan. Theory Pract. **9**, 437–457 (2008)
5. Harou, J., Pulido-Velazquez, M., Rosenberg, D.E., et al.: Hydro-economic models: concepts, design, applications, and future prospects. J. Hydrol. **375**, 627–643 (2009)
6. Pope, A.J., Gimblett, R.: Linking Bayesian and agent-based models to simulate complex social-ecological systems in semi-arid regions. Front. Environ. Sci. **3**, 55 (2015)
7. An, L.: Modeling human decisions in coupled human and natural systems: review of agent-based models. Ecol. Model. **229**, 25–36 (2012)
8. Sivapalan, M., Savenije, H.H.G., Blöschl, G.: Socio-hydrology: a new science of people and water. Hydrol. Process. **26**, 1270–1276 (2012)
9. Blair, P., Buytaert, W.: Socio-hydrological modelling: a review asking “why, what and how?”. Hydrol. Earth Syst. Sci. **20**, 443–478 (2016)

[2] http://twap-rivers.org/

[3] http://sedac.ciesin.columbia.edu/data/set/grand-v1-dams-rev01

[4] http://www.fao.org/nr/water/aquastat/main/index.stm

[5] http://web.archive.org/web/20170701012349/http://waterbase.org/download_data.html

[6] https://ccl.northwestern.edu/netlogo/

[7] www.jessrules.com/jess/

10. Barreteau, O., Bousquet, F., Attonaty, J.-M.: Role-playing games for opening the black box of multi-agent systems: method and lessons of its application to Senegal River Valley irrigated systems. J. Artif. Soc. Soc. Simul. **4**, 5 (2001)
11. Barreteau, O., Antona, M., D'Aquino, P., et al.: Our companion modelling approach. J. Artif. Soc. Soc. Simul. **6**, 1–6 (2003)
12. Khan, H.F., Yang, Y.E., Xie, H., Ringler, C.: A coupled modeling framework for sustainable watershed management in transboundary river basins. Hydrol. Earth Syst. Sci. **21**, 6275 (2017)
13. Smajgl, A., Toan, T., Nhan, D., et al.: Responding to rising sea levels in the Mekong Delta. Nat. Clim. Chang. **5**, 167 (2015)
14. Le Bars, M., Attonaty, J.-M., Pinson, S.: An Agent-Based Simulation for Water Sharing Between Different Users, pp. 211–212. ACM, New York (2002)
15. Berger, T., Birner, R., Díaz, J., et al.: Capturing the complexity of water uses and water users within a multi-agent framework. In: Integrated Assessment of Water Resources and Global Change, pp. 129–148. Springer, Dordrecht (2006)
16. Barthel, R., Janisch, S., Nickel, D., et al.: Using the multiactor-approach in G lowa-Danube to simulate decisions for the water supply sector under conditions of global climate change. Water Resour. Manag. **24**, 239 (2010)
17. Dabelko, D., Aaron, T.: Water, conflict, and cooperation. Environ. Change Secur. Project Rep. **10**, 60–66 (2004)
18. Gao, Y., Margolies, A. Transboundary water governance in the Volta River Basin. https://wikis.uit.tufts.edu/confluence/display/aquapedia/Transboundary+Water+ Governance+in+the+Volta+River+Basin (2009)
19. de Condappa, D., Chaponnière, A., Lemoalle, J.: A decision-support tool for water allocation in the Volta Basin. Water Int. **34**, 71–87 (2009)
20. UNEP-GEF Volta Project: Volta basin transboundary diagnostic analysis. UNEP/GEF/Volta/RR 4/2013 (2013)
21. Kim, K., Glaumann, K.: Transboundary water management: who does what, where. Analysing the Data in SIWI's Transboundary Water Management Database Swedish Water House, Stockholm (2012)
22. Petersen-Perlman, J.D., Veilleux, J.C., Wolf, A.T.: International water conflict and cooperation: challenges and opportunities. Water Int. **42**, 105–120 (2017)

Chapter 32
Norms in Social Simulation: Balancing Between Realism and Scalability

Cezara Pastrav and Frank Dignum

Abstract Agent-based modelling (ABM) has been used to study the dynamics of complex systems, including human societies. However, the design of such models often fails to capture one of the key features of human behavior: norms. While norms and normative behavior are extensively studied in artificial intelligence (AI), especially in the context of multi-agent systems (MAS), their approaches are often very complex and formalized, going against the prevailing discourse of ABM, which advocates keeping the models as simple as possible and pruning any unnecessary complexity. Nevertheless, norms are relevant and integral to many social contexts, and capturing their effect and dynamics often requires agents that, while not as complex as those developed for AI, are capable of sophisticated cognition. We present a normative architecture that attempts to capture the ways norms affect cognition and behavior, while at the same time being lightweight enough to be suitable for ABM use in simulations.

Keywords ABM · Norms · Normative architecture · Policy · Social simulation

32.1 Introduction

In 1990, Norway changed its fishing policy to limit fishery expansion and relieve the pressure on recently collapsed cod stocks by introducing a quota system. While this new policy had the desired effect overall, it actually had some paradoxical effects on the small-scale fisheries, which ended up expanding their fishing effort, not only

C. Pastrav (✉)
Matís ohf, Reykjavík, Iceland

Utrech Univeristy, Utrecht, The Netherlands
e-mail: cezarap@cs.umu.se

F. Dignum
Utrech Univeristy, Utrecht, The Netherlands
e-mail: F.P.M.Dignum@uu.nl

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_32

in size catches but also in time spent at sea. Where before fishers were content to fish during the autumn or the spring season, or both, now they fish year-round with unknown long-term effects on the stock. The previous fishing policy amounted to open access fishing in the case of small-scale fisheries, and fishers limited their activity based on social and economic constraints, with young fishers with larger debts fishing the most and then decreasing their activity over time as they paid back their debts, preferring to spend their time with their family or engaged in other activities. This is at odds with the assumption that the larger and more technically capable the boat, the larger the catch. It is also at odds with the economically rational assumption that fishers would try to maximize their profits. When the new quota policy was introduced, based on these erroneous assumptions about fishers' behavior, small-scale fishers became incentivized to fish more in order to meet their allotted quota or lose access to the fishery. For details on this particular case of fishing policy backfiring in unexpected ways, see [1].

This is just one of many instances where the models used to inform policy making ignore or misrepresent the behavior of the people being managed. In the case of fisheries, the focus is on mathematical stock assessment models and economic models simplified to be mathematically tractable. It is not surprising they leave out the particular complexities present in the behavior of distinct groups such as small-scale fishers. However, fisheries are complex systems with ecological, economic, and social components, and ignoring or oversimplifying any one of these elements is bound to lead to unexpected and, likely, undesirable results. It is in these situations that agent-based modeling (ABM), with its ability to handle so much more complexity of behavior and interaction than mathematical models, is suggested as a better alternative for studying complex systems. But would ABM have done much better had it been applied to the aforementioned example, where "to do better" means "to lead to better understanding of the fishers' behavior and motivations, and how policy might interfere with the already established patterns"?

Human behavior is motivated largely by social values and norms. Fishing behavior in the open access version of the Norwegian small-scale fisheries was regulated by social values and norms which dictated fishing effort should be measured based on need (debt to be repaid) and not on ability (technical capacity, experience). Spending time on shore with friends and family was more important than fishing as much as possible. Investing in better boats and equipment, too, meant better working conditions rather than the ability to catch more fish. ABM is known for preferring the simplest possible agents that appear to fit the model. While keeping the model as simple as possible works well enough in many cases, it has been criticized as under-exploitation of the possibilities afforded by the ABM approach, especially the generative aspect of it [2], and calls have been made for abandoning simplicity for more descriptive models which would allow the use of richer empirical data, including qualitative data [3]. In this case, using the simplest agents, built on the simplest assumptions, would likely result in rationally bounded utility-maximizing agents or simple ad hoc heuristics-driven agents. These agents are driven by rules directly, rather than some higher order internal motivation. The behavior we want to observe is built directly into them from the beginning.

To go beyond the limitations of such simple agents, models must employ more complex cognitive agents that are able to reason about their values, norms, and goals and modify their behavior accordingly. Unsurprisingly, such an approach involves far more time and effort, and therefore normative agent architectures are far more prevalent in fields such as multi-agent systems (MAS), where the focus is on agent cognition, than in ABM, where the focus is on the emergent behavior of the system. Nevertheless, agents capable of normative deliberation are still desirable for at least some classes of ABMs. One of them, the interaction between norms and policy, illustrated by the Norwegian small-scale fisheries example, is of particular interest. Designing policy that modulates human behavior in some desired way in a population is often complicated by the fact that usually the population is already governed by a system of norms. If the new policy being implemented clashes with well-established norms, people may choose to ignore, skirt, exploit, and generally find surprising and creative ways to not comply. When they do comply, the side effects may still be surprising and undesirable and, what's more, can vary wildly from one population to another, as is too often the case in fisheries policy.

In order to be able to implement models with agents capable of normative deliberation and behavior that are scalable to meaningful agent population sizes for simulation, we develop a lightweight architecture that captures the essential aspects of normative behavior and deliberation, while allowing future modelers to add or subtract functionality as needed for their own applications. We are aware that there is little agreement on what constitutes a norm, let alone what is essential to normative cognition, and we explain and argue for each element while highlighting the fact that our approach is by no means meant to be definitive. Because of page limitations, we only describe the individual agent cognition in this chapter, leaving the social aspects of the architecture for a future paper.

The chapter is structured as follows: Sect. 32.2 provides more information on normative agents and norms in ABM and MAS. Section 32.3 provides a description of the architecture elements and the justifications for the choice to include them over other candidates. Section 32.4 presents an example of an agent built with this architecture and its capabilities. Section 32.5 presents a discussion of further requirements for the architecture to become truly social. Section 32.6 is the conclusion.

32.2 Literature

In MAS, the focus is on agent cognition. Normative MAS agents are designed to operate in open systems of heterogeneous agents, and norms are seen as a method of helping the agents behave in ways that are individually and collectively beneficial in the absence of a central regulatory authority, much in the way norms help regulate human behavior. For a good description of what is expected from a normative agent in MAS, see [4]. A lot of effort is invested in formalizing all aspects of the architecture, but for a conceptual approach, which also describes and explains the role of many social motivators of behavior in the agent's cognition, see [5]. For a good example of a normative framework for use in MAS, see [6]. Finally, for a comparative review of MAS architectures, see [7].

As for ABM, a few illustrative examples of models that include norms include [8], which studies the emergence of norms using a game theoretical approach, and [9], which studies norm diffusion through memetic contagion. The model in [10] focuses on norm adoption based on the cognitive effort it takes to keep conforming to the current norm vs. adopting a new norm, while [11] looks at the role of norms in controlling aggressive behavior in groups. With such a diversity of approaches, it's difficult to produce a classification of approaches to modeling norms in ABM, except maybe in the broadest terms. For such a comparison of normative behavior, in terms of obedience, conformity, and compliance, see [12].

32.3 Architecture

MAS agents are endowed with complex cognitive abilities based on solid theory and thus safe from the ad hoc criticism ABM agents tend to attract. This does not mean they are suitable for ABM. In fact, it is their complexity and theoretical formalism that makes them unattractive for social simulation. First of all, the complexity of MAS normative architectures requires significant computational resources to operate, making them impossible to scale simulations to sizes demanded by ABM. Second, the ad hoc quality of ABM agents has persisted for so long for a reason: the systems being modeled are far too complex to ever be properly formalized, as opposed to the virtual environments in which MAS agents operate, which can be designed and formalized to a much higher extent. ABM modelers are forced to choose which of the characteristics of that system are the most relevant for their research questions and thus for the model and agents they're building. Some components present in MAS architectures may be irrelevant to the model, and including them can generate unwanted complexity which would make the results harder to interpret [13], not to mention the time and effort wasted parameterizing and calibrating said components.

A normative architecture suitable for ABM should, therefore, be light enough to be scalable and formalized enough to include the essential components and structures of normative behavior, while allowing the modeler to modify them or add new ones as needed. At the very least, such an architecture needs to separate motivation from behavior and include the possibility to deliberate about motivations and relate them to desired behavior. The go-to architecture for separating motivation from behavior is the BDI architectures, and normative extensions of BDI architectures (aimed at MAS) exist in literature [14, 15]. We extend the BDI template to include norms and normative deliberation, while remaining relevant to ABM.

32.3.1 Drivers of Behavior

The behavior of an agent can be said to come down to choosing what to achieve and how or choosing goals and the plans to achieve them. The choice of goals can be influenced both internally, through values, motives, and norms, and externally, through limitations imposed by the environment or externally imposed rules, such as policy rules. Internally, values and motives are usually used for selecting goals, while norms can both influence goal selection and act as goals themselves. Values and norms can also influence the type of behavior the agents selects if some (or all) actions are directly related to values, either by promoting or opposing values or by fulfilling or violating norms.

Since we can set goal priorities directly, without resorting to values or motives, and norms themselves can serve to prioritize goals, we only include goals and norms as minimal requirements for internally driving the behavior of agents in this version of the architecture. Values play an important role in deciding conflicts between norms or whether a norm is internalized or not, but these dynamics are not included for now. We leave it to the modeler to decide whether to add another level of motivation when designing their model.

32.3.1.1 Goals

Goals are simply states the agent wants to bring about. Some states are preferred over others, whether by definition or through norm influence, so there exists a preference ordering over all goals. As such, a goal is represented by one or more features of the environment and/or the internal state of the agent together with their desired values, called the condition set of the goal.

In our fishery example, fishers have goals about paying their debt and covering household costs, how much time to spend on shore versus fishing, and how much comfort to work in. A subset of fishers is also driven by the profit motive, as shown by the percentage of fishers who fish considerably more than others.

32.3.1.2 Norms

We take norms to be constraints on behavior that are socially agreed upon, rather than imposed by a central authority. Norms affect behavior by requiring certain conditions be met in order for a particular behavior to be acceptable, by dictating the effects a behavior should have or by requiring a specific behavior be performed. As such, norms can act both as conditions and goals and can introduce their own specific actions. Norms are also context specific, becoming active in certain contexts and remaining inactive otherwise.

We define a norm N as a structure [C_a, C_e, G, A, P, C], where:

- C_a is a set of pairs (a, c), where a is an action and c is a set of preconditions that must be fulfilled in order to execute action a without breaking the norm.
- C_e is a set of pairs (a, e), where a is an action and e is a set of postconditions (or effects) for a.
- G is a set of goals.
- A is a set of actions.
- P is a set of punishments for not complying with the norm.
- C is the context of the norm, defined as a set of conditions that must be true in order for the norm to apply.

The conditions follow the same structure as goal conditions. If a normative condition overlaps an existing action condition, it replaces it. Otherwise, it is added as a separate condition to the action. The goals in G replace or modify existing goals if their condition sets overlap or are added to the goal pool as new goals. At least one of C_a, C_e, G, A, C must be nonempty in order for the norm to exist. If C is absent, the norm is considered to be always active.

The norm regulating fishing behavior in the example fisheries dictates fishers should fish enough to cover their needs and no more and that their fishing effort should be proportional to the amount of debt they have. It only applies to small-scale fishers; hence, the context of the norm includes all fishers who own boats that fall under the category of "small boat."

32.3.1.3 Policy

The structure of policies is similar to that of norms. Policy can condition actions, or the effects of actions, and can impose their own actions and goals. At least one of Ca, Ce, G, A, C must non-empty in order for the policy to exist. If the context C is absent, the policy is considered to apply to everyone at all times. However, policies and norms are deliberated about differently. For details, see Sects. 3.3 and 4.5.

The policy being introduced in the second part of the example, conditions access to the fishery by requiring fishers to own quota. It also limits the amount of fishing to the amount of quota owned by the fisher, who is obliged to fish all of it, or gradually lose it and, with it, his access to the fishery.

32.3.1.4 Roles

Roles are used to determine agent type and group membership with regard to the norms and policy that apply to them, as well as which actions the agents are able to perform. As such, a role R is a structure [A, N, P], where A is the set of actions, N is the set of norms, and P is the set of policy rules that apply to the role. Agents can have more than one role and can change roles over time.

32.3.2 Actions

Actions transform environment states/internal state of the agent into other environment states/internal states. They have a set of preconditions and a set of effects, defined, similar to goals, by sets of conditions. There are three types of preconditions and effect sets: physical, normative, and policy. We define an action as a structure [f, C_{ph}, C_n, C_p, E_{ph}, E_n, E_p] where:

- C_{ph}, C_n, C_p are the sets of physical, normative, and policy preconditions, respectively.
- E_{ph}, E_n, E_p are the sets of physical, normative, and policy effects, respectively.
- f: $S \rightarrow S$, $S = \{s \mid s$ is an environment/internal state$\}$ is the function that transforms one state into another state.

The normative precondition and effects are added to an action once the corresponding norm has been adopted and are activated depending on context. The policy preconditions and effects are added once a policy comes in effect.

In our example, both normative and policy condition sets cannot be active at the same time. See Sect. 32.5 for details.

32.3.3 Deliberation

The deliberation process is split between the different manager components of the architecture (see Fig. 32.1). The managers are called in the order below:

The *norm manager* keeps track of which norms are active, decides whether the agent complies with a norm, and when a norm is internalized. An agent will comply with a norm if it doesn't conflict with its goals. In case of conflict, the norm manager weighs the cost of punishment for noncompliance versus the cost of compliance in order to decide whether to follow the norm or not. An agent always complies with an internalized norm.

The algorithm for deliberating about norms is as follows:

```
foreach(norm in Norms)
  if(norm.isActive)
    if(norm.isInternalized)comply with norm
    else foreach(goal in Goals)
        checkCompatibility(norm, goal)
      if (norm is compatible with all goals)
        comply with norm
      else
        costOfCompliance = calculateCostOfCompliance
        costOfNonCompliance = calculateCostOfNonCompliance
        if(costOfCompliance >=costOfNonCompliance)
          comply with norm
```

Norms can conflict with a goal in three ways: the norm may define an alternative to the goal that is incompatible with the preferred version of the goal, the norm may

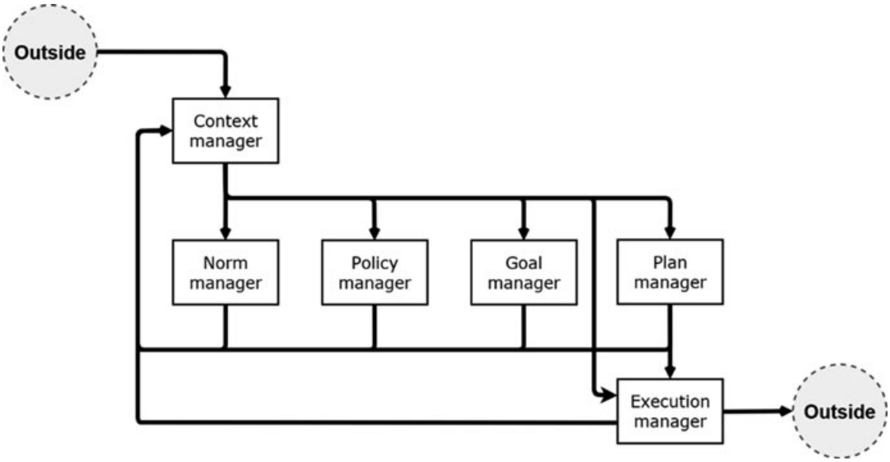


Fig. 32.1 Conceptual diagram of the architecture showing the more relevant modules and connections between them. The context manager updates the agent's beliefs and motivations, both external (environment states, policies) and internal (norms, goals). The execution manager executes the plans and actions. In between, site the four main deliberating modules (the norm, policy, goal, and plan managers)

impose conditions on actions that render plans unfeasible and the goal unreachable, and the norm may impose conditions on actions that alter plans in such a way that the goal is no longer met. For a goal to conflict with a normative goal, there must be partial or total overlap between the condition set of the goal and the condition set of the normative goal. When checking for goal compatibility, the manager checks for all three types of conflicts.

Punishments can conflict with goals in the same three ways; thus, when calculating the costs of compliance and noncompliance, the manager checks for all three types of conflicts. As a rule, goals becoming unachievable because plans have become unfeasible carry the highest cost of all three possible outcomes.

The *policy manager* keeps track of which policy is active and applicable to the agent and decides whether the agent complies with the policy by weighing the cost of punishment for noncompliance against the cost of compliance and probability of getting caught. Unlike norms, policies are not internalized, which means the temptation to not comply is always present. Its algorithm is very similar to the one of the norm manager, with additional checks for policy-norm conflicts, which are decided by comparing the noncompliance costs of the norm and the policy.

The *goal manager* keeps track of the goal queue and goal ranking. It has access to a list of active norms and policies and activates the appropriate normative and policy goals in the goal queue. Then, it selects the next active goal in queue and sends it to the plan manager. When the goal is achieved, it moves the goal to the end of the queue and selects the next one. If the plan manager cannot find a plan to achieve the current goal, the goal manager moves the goal to the end of the queue and selects the next one.

The *plan manager* keeps track of the plan library and can make new plans if needed. It has access to the list of active norms and policies the agent complies with, activates the conditions and effects sets on the relevant actions, and adds relevant norm- and policy-specific actions to the action pool. It receives a goal from the goal manager and selects appropriate plans from the library, ranks them by score, and submits the one with the highest score to the execution manager. If no plan exists, it attempts to build new plans. If successful, the new plans are scored, ranked, and the highest scoring one is sent to the execution manager. If unsuccessful, it declares the goal not achieved and requests a new goal from the goal manager.

The *execution manager* executes the plan received from the plan manager. It checks the feasibility of each action before executing it. If the action cannot be executed, it requests a new plan from the plan manager.

32.4 Example

In this section, we detail the example we presented in the paper in informal terms so far. We focus on some of the more relevant aspects of the example. The abridged version asks: how do the goals, norms, and policies influence an agent's decision about how much fishing effort to exert over a predetermined period?

32.4.1 The Goals

The fishers in the small-scale Norwegian fisheries largely fall into two categories: those who fish enough to cover their debts and living costs and those who fish to maximize their profit. Both these goals refer to the money one expects to make from fishing over a period of time; thus, we can write the two goals like this:

```
goal1 = {money > min_debt}
goal3 = {money = max_money}
```

Money obtained from fishing depends on fishing effort (measured in days at sea), the efficiency of the boat (as a function of size, gear, and the experience of the fisher), and the value of the catch. It is calculated as:

```
money = days_at_sea * efficiency * fish_value
```

Min_debt is the minimum costs the agents need to cover over a given time period and includes both the percentage of their total debt that needs to be repaid and their living costs for the same period. **Max_money** is the maximum amount of money that can be obtained through fishing over the given period. To maximize their income, fishers need to spend the maximum amount of time at sea. The income can also be increased by acquiring more efficient boats or by fishing more valuable fish, but for the purpose of this chapter, maximizing the time spent at sea is sufficient.

32.4.2 The Norm

The norm regulating this fishing behavior dictates fishers should fish enough to cover their needs and no more and that their fishing effort should be proportional to the amount of debt they have:

```
N_fishing = [G = {money = current_debt * repayment_rate},
C = {boat = small_boat}]
```

All other elements of the norm are empty.

32.4.3 The Policy

In our case, policy dictates an agent must have quota in order to fish and must fish its quota or lose access to the fishery. We can write the policy rules like this:

```
P_fishing =
[Ca = {(fishing_action, {remaining_quota > 0)},
Ce = {(fishing_action, remaining_quota -= fished_quota)},
G = {money = quota_depletion_rate * initial_quota * fish_value},
P = {fishing_action, {remaining_quota = 0}]
```

All other elements of the policy are empty.

32.4.4 Actions and Plans

In our example, the condition sets of the fishing action can be represented as:

```
fishing_action =
[C_ph = {boat != null, days_at_sea > 0},
 E_ph = {money = days_at_sea * efficiency * catch_value}],
```

where **days_at_sea** is time spent fishing measured in days and efficiency is determined by the type of boat and gear the agent has, as well as its experience in fishing. If the agent follows the norm that dictates time spent fishing is proportional to the debt, we add the corresponding normative precondition:

```
fishing_action =
 [C_ph = {boat != null, days_at_sea > 0},
    G = {money = current_debt * repayment_rate}]
```

If the quota policy comes into effect, the action becomes:

```
fishing_action =
 [C_ph = {boat != null, days_at_sea > 0},
  G = {money = quota_depletion_rate * initial_quota * fish_value},
  C_p = {remaining_quota > 0},
  E_p = {remaining_quota -= fished_quota}]
```

32.4.5 Deliberation

During the deliberation phase, the agents decide whether to comply, first with the norm, then with the policy. During deliberation, three conflicts arise and are resolved: between the norm and the profit maximization goal, between the policy and the norm, and between the policy and the profit maximization goal.

32.4.5.1 About the Norm

The norm conflicts with the profit goal because both the goal and the norm attempt to attach incompatible values to the money obtained through fishing. The norm manager identifies the conflict when checking whether the norm interferes with the goal by assessing the condition sets of the goal and the normative goal and comparing the results. The norm requires a **money** value that is less than the one required by the profit goal. Since there is no punishment attached to violating this norm, agents with profit goals choose to not comply with the norm.

The norm does not conflict with the debt repayment goal. The goal allows for any value for the **money** attribute, as long as the value covers the minimum amount required to repay the debt, and the norm ensures debt repayment by definition. Since there is no conflict between their goal and the norm, agents with the debt repayment goal decide to comply with the norm, and the debt repayment goal takes the form of the normative goal.

32.4.5.2 About the Policy

In the case of agents with the debt repayment goal, the policy conflicts with the norm because, again, they both attempt to attach incompatible values to the **money** attribute of the goal. The real-life policy assigned fishers amounts of quota that required higher fishing effort to fish in their entirety, but even if that weren't the case, the fact that debt declines over time, while quota does not, forces the conflict between the norm and policy eventually. The conflict is resolved by calculating non-compliance costs for the norm and the policy and complying with the one with the higher noncompliance cost, in this case, the policy. Interestingly, once the norm is rejected, the goal reverts to its original form, which only requires that fishing money covers the minimal debt repayment rate, without an upper limit. As such, the policy does not conflict with the original goal and is complied with. The goal takes the form of the policy goal.

In the case of agents with the profit goal, there is no norm-policy conflict since the agents do not comply with the norm. However, the policy conflicts with the profit maximization goal since the amounts of quota allocated are calculated to fall under the maximum capacity of the boats. The conflict is resolved by calculating and comparing the cost of compliance and noncompliance. Since the punishment for noncompliance is loss of access to the fishery, fishing becomes impossible and the goal becomes unachievable. As mentioned in the previous section, this carries a far greater cost than mere loss of profit. The policy is complied with and the goal takes the form of the policy goal.

32.4.5.3 About the Plan

Once the deliberation about the norm and policy is complete and the goal is set to its final form, the plan manager takes over. For simplicity's sake, we limited the plan to the necessary minimum: one action – fish. It's the plan manager's task to parameterize this one action in order to fulfill the goal. As such, it is the plan manager that, finally, gives an answer to how much effort do fishers need to exert in order to meet their goals.

In the absence of the quota policy, agents with the debt repayment goal have lower effort requirements since their goal requires less money be made from fishing, and, since they comply with the norm, their effort decreases over time. Agents with the profit maximization goal have the highest effort of all, since they refuse to comply with the norm and the quota policy is not in place to limit the amount of money they can aim to make from fishing.

When the quota policy is introduced, the agents with the debt repayment goal are forced to discard the norm and comply with the policy instead, which raises their fishing effort and maintains it at a constant level over time. Agents with the profit maximization goal are also forced to comply with the policy, and their effort level is lowered.

If the overarching goal of the policy is to lower the total effort in small-scale fisheries, then it comes down to whether the total reduction in the effort of the profit-seeking agents is greater than the total increase in the effort of agents that follow the norm. In the real-life fishery, this was not the case, but it would be interesting to calculate whether there exists a quota allocation rule that would prove successful.

32.5 Discussion and Future Work

The architecture described in this chapter aims to allow agents to reason about social norms, while remaining light enough to be scalable for social simulations. Due to space limitations, we chose to present the individual cognition aspect of the architecture, but a proper normative architecture must include social cognition as well. We will briefly touch on this aspect here.

First, the social part of norms. Norms are generally enforced socially, by other agents. This requires agents to be able to recognize normative behavior in other agents, specifically whether another agent is not complying with a norm. Once such behavior can be recognized, the agent must decide whether to punish the offending agent, which required additional deliberation capabilities.

Second, agents need to be able to communicate about norms, at least. This is especially relevant when not all agents are aware of all existing norms and need to be informed about them.

Third, a social structure connecting agents to one another needs to be in place over the agent population in order to be able to apply social pressure with regard to norms. A social network would be very suitable because it makes it easy to define norm contexts in terms of who is responsible for monitoring, and punishing whom, as well as for delineating normative contexts (work-related norms apply on the work network among agents connected by work relations, for instance).

32.6 Conclusion

We have described a normative architecture that attempts to strike a balance between the formal and rigorous normative architectures of MAS agents and the overly simplified ad hoc, but flexible, architectures of ABM agents. The goal was to build an architecture that is capable of deliberating over norms, while remaining light enough to be scalable and flexible enough that modelers can adapt it to the requirements of their own models. We also presented a real-life example of a situation where norms play an important part in the dynamics of the system (small-scale fisheries in Norway), highlighting the desirability of such architectures in ABM, and showed how it can be implemented using our proposed approach.

Acknowledgments This work was funded as part of the SAF21 project, financed under the EU Horizon 2020 Marie Skłodowska-Curie MSCA-ETN program (project 642080).

References

1. Maurstad, A.: To fish or not to fish: small-scale fishing and changing regulations of the cod fishery in northern Norway. Hum. Organ. **59**, 37–47 (2000)
2. Conte, R., Paolucci, M.: On agent-based modeling and computational social science. Front. Psychol. **5**, 668 (2014)
3. Edmonds, B., Moss, S.: From KISS to KIDS–an 'anti-simplistic' modelling approach. In: International workshop on multi-agent systems and agent-based simulation (2004)
4. Luck, M., Mahmoud, S., Meneguzzi, F., Kollingbaum, M., Norman, T.J., Criado, N., Fagundes, M.S.: Normative agents. In: Agreement Technologies, pp. 209–220. Springer Netherlands, Dordrecht (2013)
5. Dignum, F., Dignum, V., Prada, R., Jonker, C.M.: A conceptual architecture for social deliberation in multi-agent organizations. Multiagent Grid Syst. **11**(3), 147–166 (2015)
6. y López, F.L., Luck, M., d'Inverno, M.: A normative framework for agent-based systems. Comput. Math. Organ. Theory. **12**, 227–250 (2006)
7. Neumann, M.: The cognitive legacy of norm simulation. Artif. Intell. Law. **20**(4), 339–357 (2012)
8. Axelrod, R.: An evolutionary approach to norms. Am. Polit. Sci. Rev. **80**(4), 1095–1111 (1986)

9. Gatherer, D.: Identifying cases of social contagion using memetic isolation: comparison of the dynamics of a multisociety simulation with an ethnographic data set. J. Artif. Soc. Soc. Simul. **5**(4), 1–5 (2002)
10. Epstein, J.M.: Learning to be thoughtless: social norms and individual computation. Comput. Econ. **18**(1), 9–24 (2001)
11. Conte, R., Castelfranchi, C.: Understanding the functions of norms in social groups through simulation. Artif. Soc. Comput. Simul. Soc. Life. **1**, 252–267 (1995)
12. Elsenbroich, C., Xenitidou, M.: Three kinds of normative behaviour: minimal requirements for feedback models. Comput. Math. Organ. Theory. **18**(1), 113–127 (2012)
13. Squazzoni, F.: The "reign of mystery": have we missed something crucial in our experimental and computational work on social norms? In: The Complexity of Social Norms, pp. 177–188. Springer, Heidelberg (2014)
14. Broersen, J., Dastani, M., Hulstijn, J., Huang, Z., van der Torre, L.: The BOID architecture: conflicts between beliefs, obligations, intentions and desires. In: Proceedings of the fifth international conference on Autonomous agents (2001)
15. Dignum, F., Kinny, D., Sonenberg, L.: From desires, obligations and norms to goals. Cogn. Sci. Q. **2**(3–4), 407–430 (2002)

Chapter 33
Collaborating Like Professionals: Integrating NetLogo and GitHub

Nicolas Payette

Abstract We are developing an "Online Models" dialogue for NetLogo. It works like NetLogo's "Models Library" dialogue but shows every NetLogo model publicly available on GitHub instead. Our goal is to encourage the sharing of NetLogo models in a way that brings the advantages of version control systems and open-source style collaboration to the world of agent-based modelling.

Keywords Agent-based modelling · Simulation · NetLogo · Collaboration · Code sharing · Version control · GitHub

It is tempting to argue that the technical challenges that come with managing complex software artefacts are, by and large, a solved problem. Version control systems have been in common usage since the mid-1980s, and professional programmers learn, sooner or later, that ad hoc naming schemes like `myCode-v2.3_temp_Jan2018_(edits).java` are no viable alternative. Large-scale collaboration also involves social challenges. While it would be preposterous to claim that these challenges have all been overcome, the open-source software development community can boast of many successful collective endeavours. The Linux operating system and the LibreOffice suite are well-known examples, but a lot of the tools used daily by modellers are also open-source. This is notably the case for the R, Python and Julia programming languages and the Repast [8], GAMA [9, 18], MASON [14–16] and NetLogo [19, 20] agent-based modelling platforms. All of these, except R, also happen to be hosted on GitHub [3], which provides both version control (through Git) and social collaboration features.

Additional credit goes to Frank Duncan, with whom these ideas were first discussed, and to Uri Wilensky, without whom NetLogo itself would not exist.

N. Payette (✉)
School of Geography and the Environment, University of Oxford, Oxford, UK
e-mail: nicolas.payette@ouce.ox.ac.uk

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_33

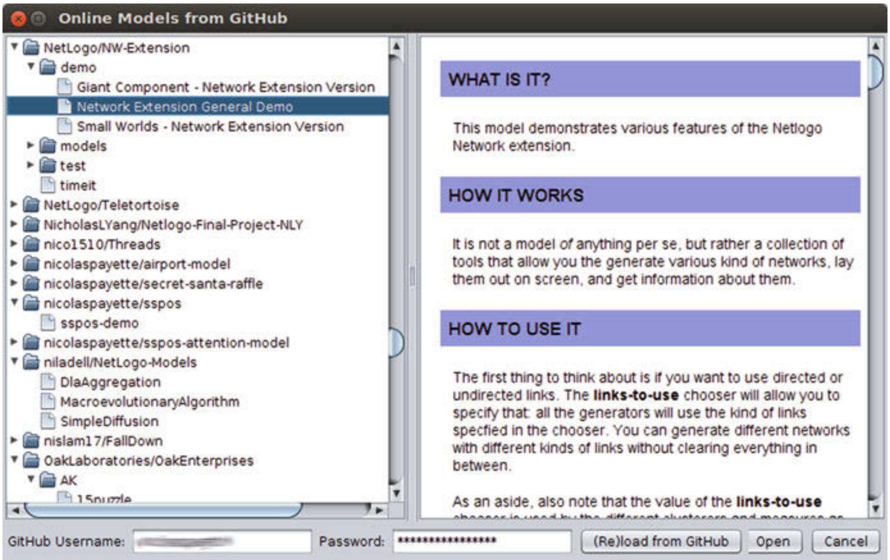


Fig. 33.1 A screenshot of the "Online Models" dialogue inside NetLogo

In spite of the benefits, academic researchers (modellers included) have not yet fully embraced the version control/open-source collaboration paradigm. About the reasons for this, we can only (but still will) speculate. Our main goal is to offer a contribution which might help take the agent-based modelling community one step closer to professional-level collaborative software development. For this, we propose to add an "Online Models" dialogue to the NetLogo user interface (Fig. 33.1). It shows every single NetLogo model made publicly available on GitHub. We will suggest future developments of this tool at the end of this paper, but will first look at obstacles to the use of version control systems in the academic community and at other possible alternatives for sharing and collaborating on NetLogo models.

33.1 Version Control Systems and Obstacles to Their Use

Simply put, a version control system is "a system that records changes to a file or set of files over time so that you can recall specific versions later". [7] This is useful for a lone programmer tracking the progress of a personal project, but it borders on essential when multiple programmers are trying to collaborate on a common project. Version control systems track what was changed when by whom. They also offer ways to resolve "merge conflicts" when multiple people tried to change the same thing at the same time.

To fully leverage the power of version control systems, the files dealt with must be plain text files: this allows the user to see exactly what was changed in each file (by herself or someone else) from one version to another and to make decisions accordingly. The need for plain text files for version control to be effective might explain part of the academic community's reluctance to embrace it: after all, the quintessential academic artefact today is arguably a paper written in Microsoft Word, a piece of software infamous for the inscrutability of its file format.[1] Users of Word somehow manage to collaborate using the "Track Changes" feature, but their proficiency with this feature is not transferable to something like an agent-based model written in NetLogo.

The learning curve that comes with the use of version control systems might be the single biggest impediment to their adoption. Some graphical user interfaces exist for them, but most version control systems still require occasional trips to the command line: a land of text and keyboard shortcuts that can seem daunting compared to icons and mouse clicks. The conceptual model associated with some version control systems also requires some getting used to. In Git, for example, the underlying model is a *directed acyclic graph* of *commits*, with *branches* and *merges*, and changes move from your *working directory* to the *repository* via a *staging area*. These concepts, unknown to most users, need to be added on top of a simple understanding of the computer's file system. Version control also requires a deeper conceptual shift: you must realize that the artefacts that you are dealing with are not just mutating over time, they carry their whole history with them. Once you have made that conceptual shift, however, the idea of working without version control becomes unbearable.

The NetLogo user experience, by virtue of its "low threshold, high ceiling" philosophy, has some elements in common with a WYSIWYG[2] tool like Word: your model is (usually) a self-contained file that includes its visual interface in addition to its documentation and code. This is both a blessing and a curse. It allows users to immerse themselves in a friendly environment with everything they need at their fingertips, without having to juggle multiple heterogeneous external tools. It does mean, however, that unless they are already familiar and comfortable with them, they tend to forfeit the power of these tools.

One of the purposes of the "Online Models" tool that we are proposing is admittedly to act as a Trojan horse: by making NetLogo models stored on GitHub easily discoverable and loadable from *within* NetLogo, we give users an incentive to integrate version control and collaboration using Git and GitHub through their whole workflow.

[1] The possibility of using version control systems with LaTeX is one of the advantages it has over Word.

[2] "What You See Is What You Get."

33.2 Alternatives for Sharing and Collaborating on Models

Model description protocols like ODD [10, 11] are useful, but not sufficient. Models should not only be described, they should be shared. It is the only way to ensure that results can be reproduced and checked. It is also the best way to increase the chances that other people use and improve the model.

The closest thing to an adopted standard for sharing models in the agent-based modelling community is the OpenABM website of the CoMSES network [1]. At the time of writing, it hosts a library of 621 models with their associated documentation, sometimes giving access to multiple versions of the models. It also offers a peer review service where models are checked for "structural completeness and fulfilling community standards". As long as the sole purpose is to make models publicly available, a site like[3] OpenABM is perfectly adequate.

Our goal, however, should not only be to share models, but also to collaborate on them. The NetLogo Modeling Commons [12, 13] goes many steps beyond simple sharing. Its features are too numerous to be exhaustively listed here, but it allows users to modify and comment on each other's models and to (crudely) compare model versions. The fact that it allows models to be uploaded from within NetLogo is also a big advantage.

There are two ways in which the Modeling Commons fall short, though. The first is that model discovery is not integrated into NetLogo. If you want to look at shared models, you have to open your browser, navigate to the website, use their search tools to find an interesting model, download it, save it locally, and then open it in NetLogo. In our "Online Models" tool, by contrast, new models are accessible in just a few clicks or keystrokes, without ever leaving the NetLogo interface. The second downside is that the Modeling Commons only work for NetLogo models. GitHub and Git, by contrast, are language agnostic and bring together most of the open-source development community. It means that models developed on GitHub are potentially discoverable by a lot more people. It also means that the proficiency acquired with Git can be extended to other project artefacts (e.g. analysis code in R, papers in LaTeX, etc.).

33.3 Future Developments

The tool, in its current state of development,[4] is pretty straightforward: it's very much like NetLogo's Models Library dialogue, except that the models are stored

[3]The "User Community Models" section of the Center for Connected Learning's website [4] plays a similar role, and non-ABM-specific scientific archiving sites [2, 5, 6] can also be used for that purpose.

[4]The tool is currently being developed in a fork of the NetLogo codebase [17], but the plan is to make it available as a plugin that can easily be installed in the official distribution of NetLogo.

online on GitHub. A browsable tree appears on the right, and the currently selected model's "Info tab" appears on the left (see, again, Fig. 33.1). If you click on the "Open" button, the model is loaded into NetLogo. GitHub users don't need to do anything out of the ordinary for their models to appear in the dialogue: as soon as the model is pushed to a public repository, it shows up in the dialogue.

We believe this sharing/discovery-enabling feature to be sufficient to justify the existence of the tool, but it is the potential for further integration with GitHub that makes it exciting. Here are a few features that could be made easily accessible from within the NetLogo interface:

– Uploading the current model to a new repository on GitHub
– Committing changes to the current model and pushing them to GitHub
– "Forking" (i.e. making a modifiable copy of) someone else's model
– Opening "pull requests" (i.e. proposing changes) for someone else's model
– Showing the differences between two versions of a model in the code tab

Taken together, these features would significantly lower the barrier to the adoption of version control for NetLogo models. Combined with the incentive of making a model easily available to all users of our tool, this might help to kickstart the adoption of open collaboration practices for agent-based modellers. In a not so distant future, perhaps our community will also be able to boast of collective achievements on the level of the greatest open-source successes.

References

1. CoMSES Net: https://www.comses.net/
2. The Dataverse Project: https://dataverse.org/
3. GitHub: https://github.com/
4. NetLogo User Community Models: http://ccl.northwestern.edu/netlogo/models/community/index.cgi
5. Open Science Framework: https://osf.io/
6. Zenodo: https://zenodo.org/
7. Chacon, S., Straub, B.: Pro Git. Apress, New York, USA (2014)
8. Collier, N.: "Repast: an extensible framework for agent simulation. Natural Resources and Environmental Issues **8**(4), 17–21 (2001)
9. Grignard, A., Taillandier, P., Gaudou, B., Vo, D.A., Huynh, N.Q., Drogoul, A.: GAMA 1.6: advancing the art of complex agent-based modeling and simulation. In: International Conference on Principles and Practice of Multi-agent Systems, pp. 117–131. Springer, Berlin, Heidelberg (2013)
10. Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S.K., Huse, G.: A standard protocol for describing individual-based and agent-based models. Ecol. Model. **198**(1–2), 115–126 (2006)
11. Grimm, V., Berger, U., DeAngelis, D.L., Polhill, J.G., Giske, J., Railsback, S.F.: The ODD protocol: a review and first update. Ecol. Model. **221**(23), 2760–2768 (2010)
12. Lerner, R.M.: Agent-Based Modeling as a Social Activity. PhD Thesis, Northwestern University (2014)

13. Lerner, R., Levy, S.T., Wilensky, U.: Connected modeling: design and analysis of the modeling commons. In: Eshet-Alkalai, Y., Caspi, A., Eden, S., Geri, N., Tal-Elhasid, E., Yair, Y. (eds.) Proceedings of the Chais Conference on Instructional Technologies Research 2010: Learning in the Technological Era, pp. 54–60 (2010)
14. Luke, S., Balan, G.C., Panait, L., Cioffi-Revilla, C., Paus, S.: MASON: a Java multi-agent simulation library. In: Proceedings of Agent 2003 Conference on Challenges in Social Simulation, vol. 9 (2003)
15. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K.: Mason: a new multi-agent simulation toolkit. In: Proceedings of the 2004 Swarmfest Workshop, Michigan, vol. 8, pp. 316–327 (2004)
16. Luke, S., Cioffi-Revilla, C., Panait, L., Sullivan, K., Balan, G.: Mason: a multiagent simulation environment. Simulation **81**(7), 517–527 (2005)
17. Payette, N.: NetLogo online models dialogue. https://github.com/nicolaspayette/NetLogo/tree/online-models (2019)
18. Taillandier, P., Vo, D.A., Amouroux, E., Drogoul, A.: GAMA: a simulation platform that integrates geographical information data, agent-based modeling and multi-scale control. In: International Conference on Principles and Practice of Multi-agent Systems, pp. 242–258. Springer, Berlin, Heidelberg (2010)
19. Tisue, S., Wilensky, U.: Netlogo: a simple environment for modeling complexity. In: International Conference on Complex Systems, Boston, vol. 21, pp. 16–21 (2004)
20. Wilensky, U.: NetLogo http://ccl.northwestern.edu/netlogo/ (1999)

Chapter 34
Kickstarting Cooperation: Experience-Weighted Attraction Learning and Norm Conformity in a Step-Level Public Goods Game

Nicolas Payette, Áron Szekely, and Giulia Andrighetto

Abstract We present a version of the experience-weighted attraction (EWA) reinforcement learning model that integrates norm conformity into its utility function that we call "EWA+Norms." We compare the behavior of this hybrid model to the standard EWA when applied to a step-level public good game in which groups of agents must reach a threshold level of cooperation to avoid the risk of catastrophe. We find that standard EWA is not sufficient for generating cooperation, but that EWA+Norms is. We aim to compare simulation results with human behavior in large-scale experiments that we are currently running.

Keywords Simulation · Public good games · Reinforcement learning · Experience-weighted attraction · Social norms · Cooperation

Recent work [12] has shown that, under some circumstances, reinforcement learning (RL) algorithms—specifically the Bush-Mosteller (BM) and Roth-Erev (RE) models—can account for human behavior in social dilemma experiments like prisoner's dilemma and public goods games (PGG), matching experimental human data as well or better than traditional "conditional cooperation" (CC) [7, 8, 13] and "moody conditional cooperation" (MCC) models [10]. Other work [14] has shown that *combining* reinforcement learning—the experience-weighted attraction (EWA) model, a more general model encompassing many other models (including RE) as special cases—with norm-based approaches like CC and MCC can match human data better than RL alone for a large-scale prisoner's dilemma experiment [9].

N. Payette (✉)
School of Geography and the Environment, University of Oxford, Oxford, UK

Á. Szekely
Collegio Carlo Alberto, Turin, Italy

Institute of Cognitive Sciences and Technologies, National Research Council of Italy, Rome, Italy

G. Andrighetto
Institute of Cognitive Sciences and Technologies, National Research Council of Italy, Rome, Italy

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_34

We are in the process of running large-scale experiments for understanding human cooperative behavior in a step-level public goods game. In this game, participants have the option of contributing part of their initial endowment of tokens to a public pot. Tokens from the pot are not given back to the participants, but if the pot reaches a known threshold, participants are protected from a catastrophe that has a fixed probability of occurring. If the catastrophe occurs without participants reaching the threshold, participants lose all their tokens for the round. Participants are organized into fixed size groups that are periodically reshuffled. Our treatments examine the behavior of agents under different probabilities of the disaster occurring, which we refer to as the risk level. We also look at what happens when the risk is modified after a given number of rounds.

Our goal is to investigate the ability of different models to match the behavior observed in this step-level PGG experiment. Ultimately, we will compare a pure RL model (EWA) and pure norm-based approaches (CC and MCC) to a hybrid model that integrates the normative component in the reward function: EWA+Norms. Decades of theoretical and empirical work, both observational and experimental, show that when deciding whether to cooperate, humans do not always make choices that maximize their individual material payoffs: they also care about complying with the norms of their group. Social norms are informal rules that prescribe what individuals ought or ought not to do and are typically enforced through informal sanctions, such as ostracism, negative gossip, shame, or disapproval [1, 5, 6]. In order to account for this twofold motivation driving human cooperative behavior, we extend the EWA algorithm—which is specified exclusively by individual considerations—to incorporate normative considerations. Our hypothesis is that this hybrid model will outperform the pure approaches.

This abstract reports on preliminary work being done on the simulations. The empirical data is not yet available for matching the simulation results, and we have so far concentrated our efforts on exploring the properties of the EWA+Norms algorithm. The key result so far is that even in high-risk settings, where incentives to cooperate are very high, EWA by itself is unable to generate cooperation, but giving our agents a bit of normative motivation is enough to kickstart cooperation.

34.1 The EWA+Norms Model

First proposed by Camerer and Ho, the EWA model has been fairly successful at replicating human behavior for a variety of games [2, 3]. The core idea of EWA (and of reinforcement learning models in general) is that agents maintain a set of "attraction" values for each possible action that they can take in a given situation. Actions that lead to positive outcomes get reinforced and are thus more likely to be chosen in subsequent rounds. Equation (34.1) shows how the attraction for an action j is updated for agent i at time t. The basic mechanism is that, following each round of the game, agents use the payoff they got to update the attraction of the action they chose, but they also use the payoffs they *would* have gotten, given the actions taken by other players, to update the attractions of the actions they did *not* take (see [2, 3] for details).

$$\underbrace{A_i^j(t)}_{\text{new attraction}} = \frac{\overbrace{\phi}^{\text{decay rate}} \cdot \overbrace{N(t-1)}^{\text{observation equivalents}} \cdot \overbrace{A_i^j(t-1)}^{\text{previous attraction}} + \overbrace{[\delta + (1-\delta) \cdot I(s_i^j, s_i(t))]}^{\text{weight of hypothetical actions vs. actual action}} \cdot \overbrace{\pi_i(s_i^j, s_{-i}(t))}^{\text{expected payoff given actions of others}}}{\phi \cdot N(t-1) + 1} \tag{34.1}$$

The generality of the EWA model comes at a price; however, it has a lot of free parameters, which makes it vulnerable to overfitting. To remedy this situation, Ho and Camerer [11] developed a "self-tuning" version of the model, where all but one of the parameters are either fixed or "set by a function which depends on observed experience." This is the version we use in our simulations.

As in [14], we explicitly incorporate people's sensitivity to social norms, their norm psychology [4], into EWA. To account for conditional cooperation, agents' decisions to comply with norms are positively affected by the share of peers that complied with them. To account for moody conditional cooperation, agents are also more likely to comply with norms and to be influenced by the compliance of others if they obeyed norms themselves in the previous round. To capture these ideas, we replace the payoff in Eq. 34.1 by a utility function adding a normative component to that payoff, where the norm to be followed is the norm of cooperation. In Eq. 34.2, we let $c = 1$ if agent i cooperated in the previous round and $c = 0$ otherwise, and we let o be the proportion of other agents that cooperated. In the context of our step-level PGG, "cooperating" means contributing at least the amount needed for the group to reach the threshold if all players contribute in equal parts. The crucial parameter in Eq. 34.2, however, is w_n, a unitless, unbounded quantity constituting the "normative weight": the importance given to social norms by agents when making their decisions. For high values of w_n, the potential monetary payoffs, while never completely ignored, are dwarfed by the agents' tendency to follow social norms of cooperation. Setting $w_n = 0$ gives us back the pure EWA model.

$$u_i = \overbrace{\pi_i}^{\text{monetary payoff}} + \overbrace{w_n}^{\text{normative weigth}} \cdot \frac{\overbrace{c}^{\text{whether or not } i \text{ cooperated}} + \overbrace{o}^{\text{proportion of other agents that cooperated}} + \overbrace{oc}^{\text{interaction}}}{3} \tag{34.2}$$

34.2 Preliminary Results and Discussion

In the results we report here, we simulate four groups of six agents each. We use an initial endowment of 100 tokens and a threshold of 300 tokens. Since each group has 6 participants, each of them must contribute 50 tokens for the group to reach the threshold. Each simulation lasts 28 rounds. Groups are reshuffled every seven rounds, except for the last seven rounds where they are reshuffled each round.

We have four treatments with different risks of the catastrophe occurring if participants do not reach the amount that the group has to contribute in order to be protected from it. In two of these treatments, the risk is held constant throughout the simulation, with a risk of 0.25 and a risk of 0.75, respectively. In the other two treatments, the risk changes mid-simulation (at round 14). In one of them, the risk goes from low to high (0.25 → 0.75). In the other, it goes from high to low (0.75 → 0.25). For each of the four treatments, we ran 100 simulations for each value of $w_n \in [0..100]$, for a total of 40,400 runs. The results are summarized in Fig. 34.1, where we show the mean contribution of players over time and normative weights for different risk levels.

The first thing to notice is that, in all four treatments, a normative weight of zero leads to defection across the board. In the case where the risk is 0.25, this is rational behavior: the expected individual payoff, when not contributing, is 75 tokens—better than the best possible payoff when contributing 50 which, assuming that everyone else also contributes or the catastrophe does not occur, is 50 tokens (see Fig. 34.2). With a risk of 0.75, however, everyone would be better off cooperating since the expected payoff of defection is 0.25 tokens while the payoff when everyone is cooperating is 50 tokens. This makes full cooperation a Nash equilibrium, since no one would be better off by defecting, but unlike in the PGG and PD games, in which RL algorithms have been shown to achieve cooperation, EWA fails to

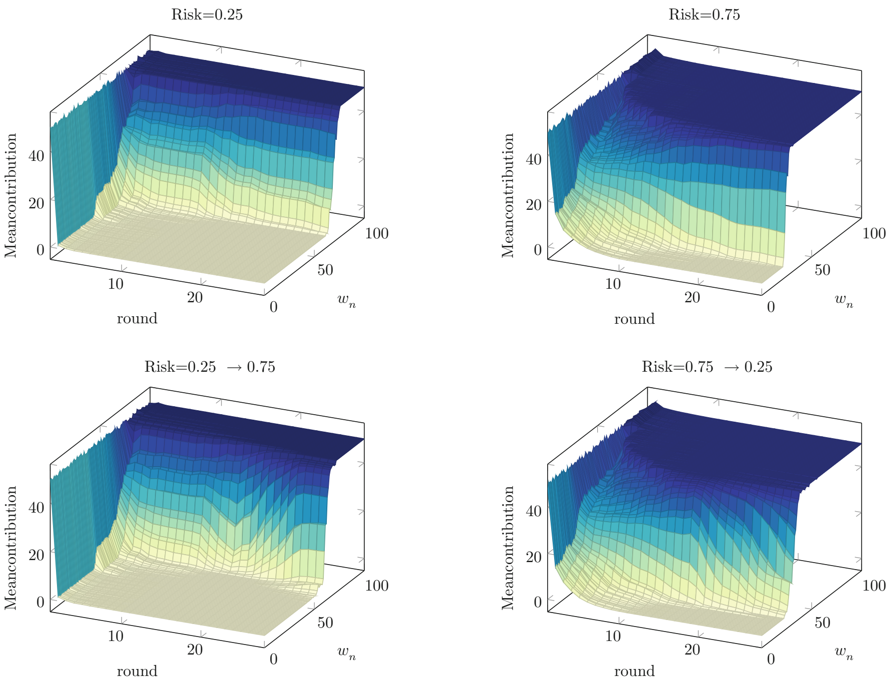


Fig. 34.1 Mean contribution over time and normative weight for different risk levels

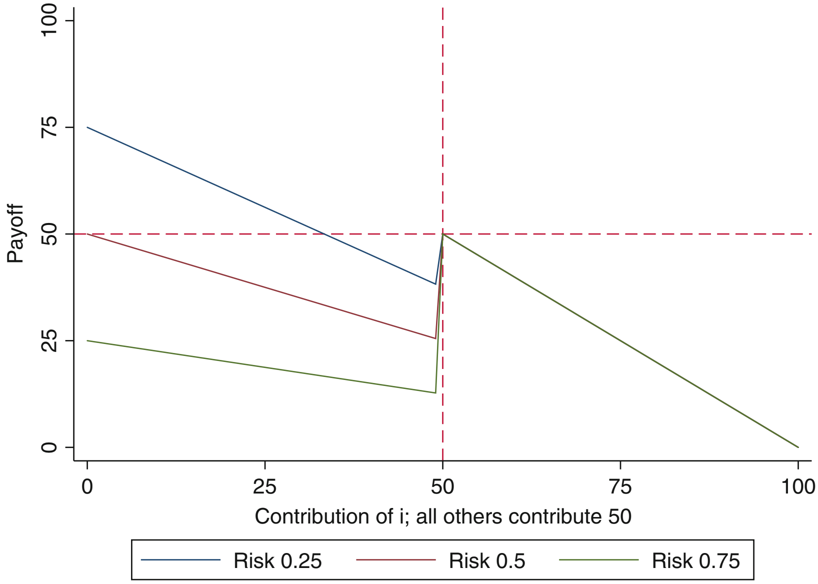


Fig. 34.2 Payoff at different contribution levels when all other players contribute 50

reach it. Giving our players some normative motivation, however, is sufficient to kickstart cooperation: at $w_n = 22$, some players start to maintain cooperation until the end of the run; at $w_n \geq 28$, all players do. With a risk of 0.25, agents don't start cooperating unless $w_n \geq 62$ and don't fully cooperate until $w_n \geq 76$. Assuming that human players would cooperate when risk is 0.75, but not when risk is 0.25, this gives us a range of plausible values for w_n. We also looked at what happens when the risk level changes mid-run. We find a limited range of w_n values where agents are able to adapt their behavior to that change ($53 \leq w_n \leq 71$ when risk goes up, $23 \leq w_n \leq 37$ when risk goes down). It will also be interesting to see if the behavior that we will observe in experiments falls into that range.

At the present stage, we can only explore the abstract behavior of our EWA+Norms model, but we will soon have empirical data to compare it too. How well will the model be able to fit the data? How does this model compare to other models of human cooperative behavior? Which value of w_n will give us the best fit? Can we assume that all human players pay the same attention to social norms, or will we find that players can be clustered into different types? Answers to these questions, whether positive or negative, should be obtained in the near future.

References

1. Bicchieri, C.: The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press, New York, USA (2005)

2. Camerer, C., Ho, T.H.: Experience-weighted attraction learning in coordination games: probability rules, heterogeneity, and time-variation. J. Math. Psychol. **42**(2), 305–326 (1998). https://doi.org/10.1006/jmps.1998.1217
3. Camerer, C., Ho, T.H.: Experienced-weighted attraction learning in normal form games. Econometrica **67**(4), 827–874 (1999)
4. Chudek, M., Henrich, J.: Culture–gene coevolution, norm-psychology and the emergence of human prosociality. Trends Cogn. Sci. **15**(5), 218–226 (2011)
5. Conte, R., Andrighetto, G., Campennl, M.: Minding Norms: Mechanisms and Dynamics of Social Order in Agent Societies. OUP, New York, USA (2013)
6. Elster, J.: Nuts and Bolts for the Social Sciences. Cambridge University Press, Cambridge, UK (1989)
7. Fehr, E., Gintis, H.: Human motivation and social cooperation: experimental and analytical foundations. Ann. Rev. Sociol. **33**(1), 43–64 (2007). https://doi.org/10.1146/annurev.soc.33.040406.131812
8. Fischbacher, U., Gächter, S., Fehr, E.: Are people conditionally cooperative? Evidence from a public goods experiment. Econ. Lett. **71**(3), 397–404 (2001)
9. Gracia-Lázaro, C., Ferrer, A., Ruiz, G., Tarancón, A., Cuesta, J.A., Sánchez, A., Moreno, Y.: Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma. Proc. Natl. Acad. Sci. **109**(32), 12922–12926 (2012). https://doi.org/10.1073/pnas.1206681109
10. Grujić, J., Fosco, C., Araujo, L., Cuesta, J.A., Sánchez, A.: Social experiments in the mesoscale: humans playing a spatial prisoner's dilemma. PLOS ONE **5**(11), e13749 (2010). https://doi.org/10.1371/journal.pone.0013749
11. Ho, T.H., Camerer, C.F., Chong, J.K.: Self-tuning experience weighted attraction learning in games. J. Econ. Theory **133**(1), 177–198 (2007). https://doi.org/10.1016/j.jet.2005.12.008
12. Horita, Y., Takezawa, M., Inukai, K., Kita, T., Masuda, N.: Reinforcement learning accounts for moody conditional cooperation behavior: experimental results. Sci. Rep. **7**, 39275 (2017). https://doi.org/10.1038/srep39275
13. Keser, C., van Winden, F.: Conditional cooperation and voluntary contributions to public goods. Scand. J. Econ. **102**(1), 23–39 (2000)
14. Realpe-Gómez, J., Andrighetto, G., Nardin, G., Montoya, J.A.: Balancing selfishness and norm conformity can explain human behavior in large-scale prisoner's dilemma games and can poise human groups near criticality. Phys. Rev. E **97**(4), 042321 (2018)

Chapter 35
Using Agent-Based Simulation to Understand the Role of Values in Policy-Making

Antoni Perello-Moragues and Pablo Noriega

Abstract We propose to explore the role of values in policy-making and the use of ABS for elucidating this role. In this paper we outline a conceptual framework for value-driven modelling of public policies and illustrate it with an agent-based simulation of irrigation practices.

Keywords Agent-based simulation · Policy-making · Policy values

35.1 Introduction

It is generally acknowledged that policy-making is about achieving a better state of the world, and, consequently, at design time, it implies that policy-makers make choices based on values [14, 45]. However, the actual effects of those choices are difficult to assess, since the real world is complex [25]. Moreover, policy-makers need to address trade-offs between the conflicting interests of the several stakeholders. We postulate that one way to deal with this complexity is to elucidate how values are involved in those decision-making processes.

For this purpose, we are following a threefold strategy: (i) first, we develop a theoretical framework that articulates the interplay between the activities involved in the policy-making cycle and the value-based choices of the main stakeholders; (ii) second, we propose to use agent-based simulation (ABS) to visualise the relationships between policy goals and instruments, on one side, and the behaviour

A. Perello-Moragues (✉)
IIIA-CSIC, Barcelona, Spain

Aqualia, Madrid, Spain

Universitat Autònoma de Barcelona, Bellaterra, Spain
e-mail: tperello@iiia.csic.es; antonio.perello@fcc.es

P. Noriega (✉)
IIIA-CSIC, Barcelona, Spain
e-mail: pablo@iiia.csic.es

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_35

of those agents for whom the policy is intended, on the other; (iii) third, we focus our attention on a specific policy domain—the use of water in agriculture—in order to draw inspiration from realistic examples, have access to empirical data and expert advice, and develop guidelines for a wider application of our proposal.

In this paper we show how ABS may be used to explore the role of values in the agenda-setting stage of the policy-making cycle. More specifically, we outline our conceptual framework in Sect. 35.3 and its background in Sect. 35.2. In Sect. 35.4 we discuss a model that illustrates the gist of our proposal. The model is based on actual agricultural data and practices, but we present a simplified version to illustrate the interplay between two policies—based on different values—for regulating irrigation practices in a community of farmers whose individual decisions are guided by different value sets. After the discussion of this example, we identify some key challenges and sketch our future research plan.

35.2 Background and State of the Art

35.2.1 Values and Behaviour

Values are at the core of decision-making, motivations, preferences, and attitudes—although they are not the same concepts [18, 34]. Schwartz et al. [41] defined values as "concepts or beliefs, about desirable end *states* or *behaviours*, that transcend specific situations, guide selection or evaluation of behaviour and events, and are ordered by relative importance".

Rohan [37] pointed out the definitional multifariousness of the "value" construct. Nevertheless, there is a consensus that values play an active role in the intentional human behaviour and decision-making—regardless of whether the reasoning about them is conscious or not [34].

This leads to postulate that individuals hold different *rationalities* [42]: decisions are made because they serve the values of an actor *as the actor sees them*. Besides values, they use other constructs, such as mental models, that are used to understand and interact with the environment [26], that provide explanations and inferences about diverse phenomena [10], and often are incomplete and biased [6, 17, 26]. We assume, therefore, that agents hold mind-frames, containing values and other constructs that support decision-making. They are often collectively shared, like other socially developed constructs [17, 18, 28].

There is no consensus on the categories of values. We identify, as fundamental categories, (i) *individual values* and (ii) *social values* (see [38]). *Individual values* refer to those values towards satisfying needs and self-esteem [39]. In contrast, *social values* are values of society at large, concerning public interest or contribution to well-being, that emerge from the society or social group [28] and that would include also desirable properties with regard to governance [49]. However, this emergence is not trivial. When referring to social values of a single agent, one may assume that such emergence is based in the agent's consideration of the other agents' minds in the group (even if this social value is not consensual) [15]. In short, individuals have beliefs about what defines the well-being of a society altogether and how public affairs should be governed in order to achieve a good social outcome (i.e. many individuals *live well*), and these may be incompatible with their individual values.

35.2.2 Values in Policy-Making

In broad terms, we understand policy-making as a process through which a group of agents, that we call *policy-makers*, design, enact, and evaluate a set of instruments to govern the activity of other agents, that we call *policy-subjects*, within some domain of activity [45]. A policy is devised in order to *govern* the activity of policy-subjects towards a state of the world that is deemed desirable by policy-makers as well as other relevant stakeholders [29]. Hence, governance is achieved through *means* like norms, incentives, and programmes, towards *ends* or intended outcomes of the policy. These policy ends are usually represented by some performance indicators or metrics that serve to assess whether the policy means are being successful [24].

According to Stewart [45], "policy design has a value-based component because the ways we attempt to change or influence behaviour depend on, in turn, beliefs about the reasons for that behaviour". In other words, the choice of policy means and ends reflects the mind-frames of policy-makers [14, 16, 45]. The adoption of a policy, on the other hand, depends on the decisions that policy-subjects make, thereby involving their mind-frames and, in particular, their values [13, 27, 29]. Thus, we assume that policy values are instantiated as means and ends according to their makers' mind-frames [45, 46].

35.2.3 Reasoning with Values

We adopt a consequentialist approach, that is, we reason about values through their effects: choose *means* because of the effects they achieve and *ends* through indicators that reflect those values (see [35]).

Moreover, we recognise that values may be *prima facie* incommensurable and thus lead to value conflicts [35]. Trade-offs may be based on the assumption of an ordering of values by relative importance invoking some form of "satisficing" combinations [44].

Finally, we assume that policy-makers reason on values by means of argumentation. Some works have approached practical reasoning in argumentation frameworks (for instance, [3, 48]). Not to mention the role of values in design (as in policy-making and computational modelling [5, 30, 35]).

35.2.4 Simulation in Policy-Making

Because of the complexity of policy-making, the use of simulation in policy design has been advocated to reproduce the dynamics of an artificial system so as to observe their behaviour and afterwards draw inferences and conclusions [21]. It enables to explore alternative policies without committing resources and disturbing the real social system, which produces useful evidence in order to identify successful and counterproductive pathways [4, 25].

Basically, an agent-based model (ABM) generates an "artificial society" from empirical knowledge that will be deployed for computational simulation [20, 22]. Ideally, these models capture the variety of decision-makers that interact within the system, which is particularly useful to model agents with distinct mind-frames as aforementioned.

Agent-based simulation has been used to study diverse policy domains: water resources management [8], land-use changes [33], agriculture [7, 40]; environmental programmes enrolment [43], and R&D policy [1], among others.

35.2.5 Water as a Policy Domain

Water, as an essential resource, has an impact across the whole socio-economic and ecological system. Not surprisingly, water is at the heart of several social and ecological conflicts—which are aggravating due to the global change [12]—and thus involves a rich repertoire of challenges in value-driven decision-making, argumentation, negotiation, collective agreement, and, in general, value-driven policy-making. Consequently, water management implies value-charged decisions, trade-offs, and ethical judgements of many sorts [9, 11, 16, 23, 32, 47, 50]. We believe that ABS is well fitted to address those challenges.

35.3 An Outline of a Conceptual Framework

We presume that values play a substantial role in policy-making. To this end, we propose to develop a conceptual framework to represent *policy-making as a value-driven social coordination process*. The core of this proposal is outlined in Fig. 35.1.

In broad terms, we postulate that policy-making involves *collective processes* where policy-maker agents *institute means and ends* in order to *influence the behaviour* of policy-subjects so as to achieve a *better state* of a relevant part of the world. We say that these processes are *value-driven* because the choice and assessment of means and ends reflect the values of policy-makers, while the behaviour of policy-subjects responds to their private values.

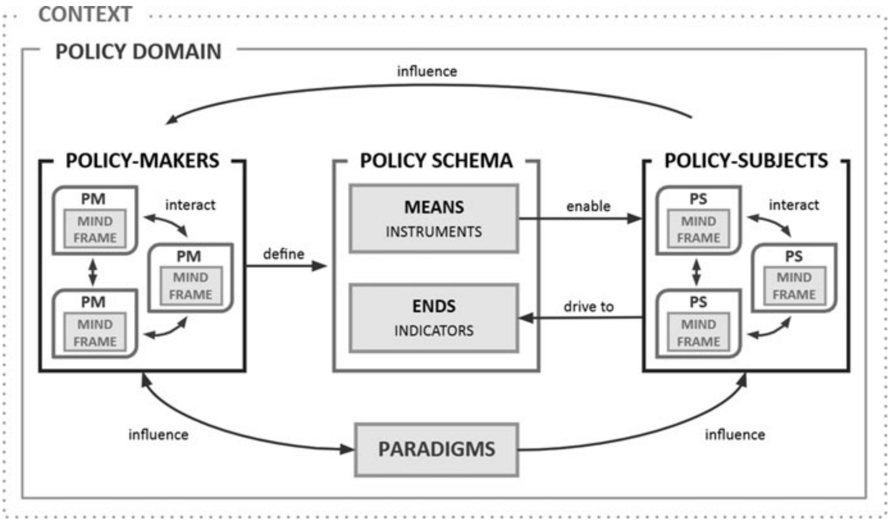


Fig. 35.1 Distinctive features of policy-making as a value-driven socio-cognitive system

We can be more precise:

1. **Policy-making processes are a subclass of socio-cognitive technical systems** (SCTS) [31]. Consequently, a policy-making process $\mathcal{P}$ has the following features:

   s.1 $\mathcal{P}$ is situated in a physical/socio-economic context, and it organises and refers to activity and entities of a limited relevant fragment of that context: the *policy domain*.

   s.2 Involves a class of stakeholders that contains at least two distinctive agent roles: *policy-makers* and *policy subjects*.

   s.3 Agents behave according to their own *mind-frames*, which include personal values among other cognitive constructs.

   s.4 A *policy domain ontology* that is the same for all stakeholders

   s.5 An observable *shared state of the world* that is altered by events and actions.

   s.6 Actions are *conditioned* by physical and normative constraints, but they may be further constrained by norms and conventions whose compliance and enforcement is determined and applied by stakeholders.

   s.7 Agent interactions are organised in policy-making *action arenas* [36] ("scenes" [19]). We distinguish three arenas (that reflect the "policy-making cycle"):

      (i) *Policy definition:* policy-maker factions negotiate their preferred policy means and policy ends and agree on their instrumentation (actions, norms, messages) and assessment (metrics, indicators).

      (ii) *Policy enactment:* policy is proclaimed, policy-subjects perform afforded actions in the world of interest subject to the policy-related norms and alter the state of the world. Policy-subjects may act individually or collectively. Each agent has its own mind-frame that conditions its behaviour. In particular, individuals have their own values and metrics associated to their individual actions and the social outcome.

(iii) *Policy evolution:* policy-subjects may become aware of desirable changes in policy ends and means and negotiate among themselves and eventually with policy-makers.

2. **Policy-making processes are value-driven** (see Sect. 35.2):

v.1 *Assumptions about values:*
   - We hold a *consequentialist* view of values. This entails that values are meant as their effects and thus may be *projected* by individuals or groups onto policy means (i.e. a norm ϕ or an action μ that promotes value α and demotes value β) and policy ends (i.e. world state σ is better than state τ according to value α).
   - We distinguish between *private values* that are held by individual agents and collective agents and are involved in their decision-making and *public values* that are involved in the assessment of the "goodness" of actions, events, facts, governance means, and states of the world.
   - For both types above, we distinguish between *individual* and *social* values. The first refer to individual agents, and the second are attributed to social groups.
   - We do not require that values are commensurable and do not commit to the existence of forms of aggregating *individual* and *social values*.

v.2 *Policy schema:* is the explicit expression of the use of values in the way a policy will be instrumented and assessed. We distinguish two main constructs:
   - *Policy means* are meant to foster activity of policy subjects towards the policy objectives. They essentially define, regulate, and motivate a set of *institutionally affordable* actions that are supported by *instruments* such as incentives, norms, and persuasion discourses.
   - *Policy ends* define desirable future states intended to be achieved by the policy and are specified through a set of *metrics* and *indicators* to measure the evolving state of the social space.

v.3 *Policy evaluation:* policy-makers will draw on the policy schema to assess the success of a policy; however, agents may formulate additional ends and metrics (possibly kept private) and using the schema metrics and their own obtain a different assessment of the outcomes.

3. **Paradigms and mind-frames:**

p.1 Paradigms [13, 32] consist of social values, norms, and practices, as well as a shared ontology that recognises facts and actions and allows for intelligible representations of the world. A paradigm is somehow assumed by society and its members and is thus reflected in individual mind-frames and collective worldviews.

p.2 Paradigms are instantiated as *policy paradigms* that constitute legitimised worldviews adopted by policy-makers and policy-subjects. They prescribe means and ends seen as suitable for specific issues, as well as proscribe others.

p.3 In the policy definition and evolution arenas, policy-maker factions and organised policy subjects strive to steer public policies according to their mind-frames. Factions may try to redefine paradigms.

4. **Uses of the framework:**

u.1 To **model** a policy-making process in a given domain (we refer to this as the *real-world model*). Roughly, it serves for structuring the elements of the SCTS: policy-maker factions, policy-subjects with their values and frames, policy domain governance infrastructure, and policy schemata. It may have descriptive (understand the policy domain) or prescriptive roles (support the implementation of a policy schema).

u.2 A **specification for an agent-based simulation** that may be used to understand a policy domain, to explore potential interventions, and to argue about policy schemata. Although all arenas may be part of the implementation, the policy enactment arena is the one where ABS has proven most useful. Figure 35.2 depicts the core elements of the conceptual framework involved in the simulation of enactments.

u.3 The implementation of the model may provide the grounds for a **policy-support systems** that may be used in the different stages of the policy cycle. Note that once the policy is active, the implemented model may still work for monitoring actual outcomes.

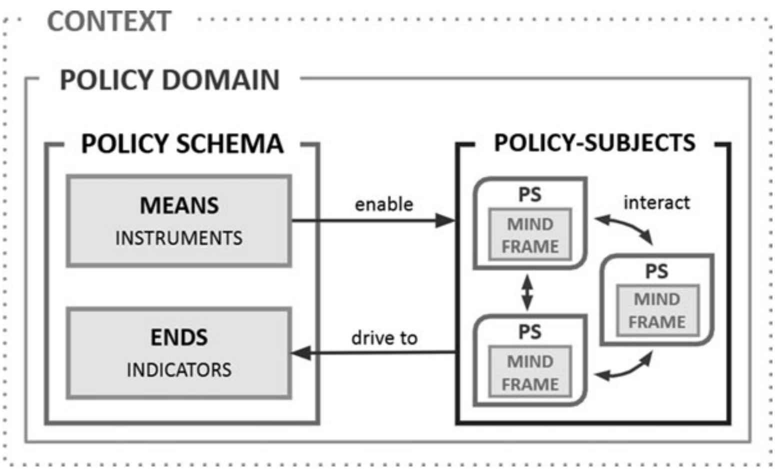


Fig. 35.2 Simulation components in the policy-enactment arena

35.4 An Illustrative Exercise

We model a crude agriculture environment and simulate the enactment of two simplistic policies with policy subjects that conform to different value profiles. The point of the exercise is to exemplify the notions of policy schema, means, and ends and to illustrate the interplay of policy-maker and subject agent values.

Model The model represents a community of farmers who draw water from a watercourse, use it to irrigate a crop, and sell their production for profit. Rain provides both farmers and the watercourse with resources.The purpose of the model is to test policies (introduced as norms and actions). That is, to observe their effects on the socio-economic environment and their acceptance by stakeholders.

Scales The model simulates three decades of activity through discrete time steps. Each month it rains, water flows, and farmers irrigate. Each year, farmers harvest and sell their production and evaluate their happiness.

The spatial scale of the model represents a watershed where farmers share a water source. The land is represented with square patches. Patches have coefficients that are used to calculate the water runoff and the percolation. Precipitation and evapotranspiration are constant, uniform across the watershed, and distributed over the year by months (e.g. in summer, it rains less and evapotranspiration is higher). We model the watershed as two inclined rectangular planes that discharge into a water channel in the middle. The whole watershed discharges into a point where a flowmeter is placed.

Policy-makers We have two factions. One supports policy P_1; the other supports P_2 (see below). Both factions agree that irrigation agriculture is desirable for the public interest—which is the shared *paradigm*.

Policy-subjects Farmers are characterised by the following attributes: (i) *mind-frame*; (ii) *farm size*; (iii) *money*; and (iv) *crop water requirement*. Also, farmers are able to (a) *grow crops*; (b) *expand* their farm; and (c) *leave* the farm. With regard to action (a) farmers demand water to complement the water required by the crop they grow—calculated as the reference evapotranspiration multiplied by a crop coefficient [2]—that is not supplied by the soil of their location.

Farmers' *mind-frame* can be either *environmentalist* or *productivist*. Farmers that subscribe to productivist values focus on their *income* and the satisfaction of their irrigation demands. In contrast, *environmentalist* farmers focus on the environment (i.e. on the volume of water in the watercourse) and, to a lesser extent, on their *income*. We assume that policy-subjects have private metrics.

Mind-frames (and values) determine the farmers' *happiness*, which is modelled as a dichotomous state. *Happy* farmers irrigate complying with applicable water constraints. However, *unhappy* farmers adjust their behaviour in order to improve their chances to be happy: *productivist* farmers ignore water constraints according to a probability, as long as they have less than a minimum amount of *money*; moreover, when they have enough *money*, they invest to increase their *farm size*, insofar as it is allowed. In contrast, *environmentalist* farmers always comply with constraints, and when they have enough *money*, their water withdrawal is only half of their actual demand.

Policy schema Policies include as *means* the action that farmers can irrigate their crops and sell their production for their own profit.

Policy P_1 embraces values like *productivity* and *wealth*. It focuses on the rural development of the basin, understood as the growth of farm industries and the wealth they generate (*end*). It considers that promoted values may be measured as the average economic resources of farmers (*metrics*)—although there are other options, for instance, the total farming area of the basin. In parallel, the policy posits that the desired state is achieved when farmers can irrigate with low restrictions and they are able to expand their farms at will (*means*).

Policy P_2 holds environmentalist values like *conservationalism* and *fairness*. It aims at keeping the watershed in good environmental conditions (*end*). It establishes that the flow of the watercourse should be regularly monitored (*metrics*), and they support this statement with scientific studies that define an environmental threshold. Consequently, farmers can irrigate, but they cannot expand their farms (*means*). Also, they cannot extract more than a certain amount of resources. Additionally, their turn for irrigation will be determined according to their money amount, and not by their location, as it was by default (*instruments*).

35.4.1 Simulations

Experiment 1 focuses on the comparison between policy (factional) values.

Both policies are tested on an environment inhabited by a mixed population of policy-subjects. Let us assume that the value profile of the population is 50% *productivist* and 50% *environmentalist*. Policy values of P_1 and P_2 are input into the simulation model as shown in Table 35.1 (three instruments and one metric for each policy).

The evolution of the simulated effects of both policies, according to the established metrics, are plotted in Figs. 35.3 and 35.4. As expected, the watercourse flow under P_1 has more acute seasonal variation, falling under the environmental

Table 35.1 Comparison between policy values (as input for the simulation)

Input	P_1	P_2
Water constraint[a] (m^3/ha/year)	10,000	1,000
Farm expansion	Enabled	Disabled
Turn system (based on)	Distance	Money
Main metrics	Money (per capita)	Flow (m^3/month)

[a]Distributed equally per month

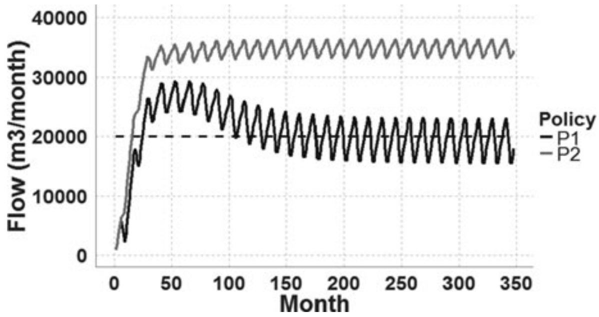


Fig. 35.3 Monthly flow evolution under both policies

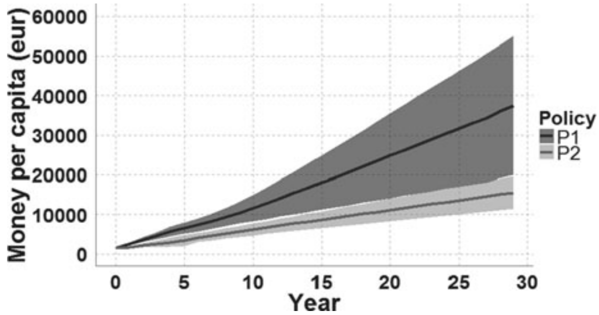


Fig. 35.4 Money (average and standard deviation bands) evolution under both policies

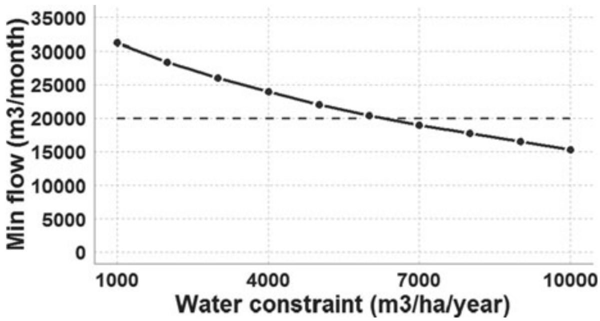


Fig. 35.5 Minimum monthly flow of the year according to the water constraint

recommendation repeatedly. In contrast, the average of the money of farmers rises faster under P_1, although the deviation is lower under P_2.

These simulation results would be used to negotiate *instruments* in order to reach an *end* that satisfies both factions in a policy definition arena. For example, one may assume that both factions reach an agreement with regard to the other *means*, and therefore the negotiated policy (i) must enable the expansion of farms and (ii) must base the irrigation turns on farmer's worth. They focus on the long-term effects (in this case, after 30 years) to negotiate water constraints (Figs. 35.5 and 35.6). Observing the effects of policies, policy-makers could negotiate for a suitable water constraint, which would lead to the emergence of a consensual *social value*, as defined in Sect. 35.2.

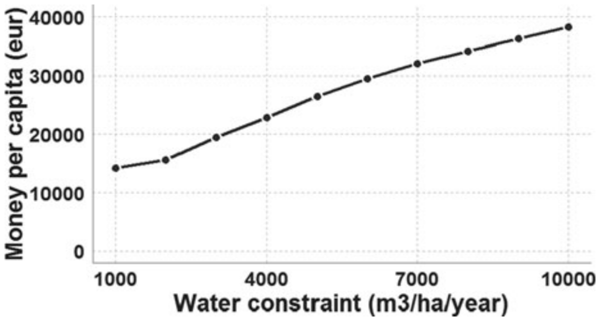


Fig. 35.6 Money (average) according to the water constraint

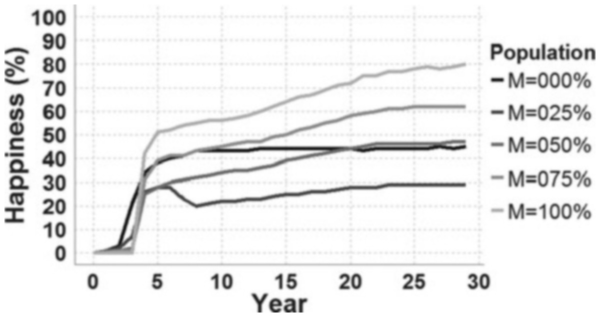


Fig. 35.7 Happiness (%) with respect to mind-frames mix, under policy P_1

Experiment 2 contrasts the values of policy-makers and those of policy-subjects. That is, given a policy, how policy-subjects' support varies when altering their value profiles.

In this case, we do not care about the effects on the world, but rather the acceptance of policy-subjects, measured as the percentage of happy farmers. With P_1 as input (Table 35.1), we change the population of *productivist* (M) and *environmentalist* ($E = 100 - M$) farmers to observe the outcome.

Not surprisingly, the more *productivist* the farmers are, the happier (Fig. 35.7), since the policy values of P_1 match productivist values. However, the point of these experiments is to gauge acceptance in populations with mixed mind-frames. Curiously, when the population is completely *environmentalist* ($M = 0\,\%$), the amount of happy farmers is nearly 50%. The reason is that farmers do not disturb the ecosystem in spite of having a low water restriction—remember that they irrigate only a fraction as long as they have a certain amount of money. However, whenever there are *productivist* farmers ($M > 0\,\%$), they use more resources—and even more when they increase their farms—driving *environmentalist* farmers unhappy.

We also compare farmers' acceptance (*happiness*) with water constraints for various population profiles (Fig. 35.8). Predictably, opposite population profiles have curves with slopes of opposite sign. When the population subscribes completely to environmentalist values, the constraint that leads to full happiness is the one promoted by P_2. On the contrary, a productivist population is happier with the lower constraints of P_1. Nevertheless, the latter does not achieve full happiness, because when all farmers increase their farms, the aggregate demand depletes the ecosystem.

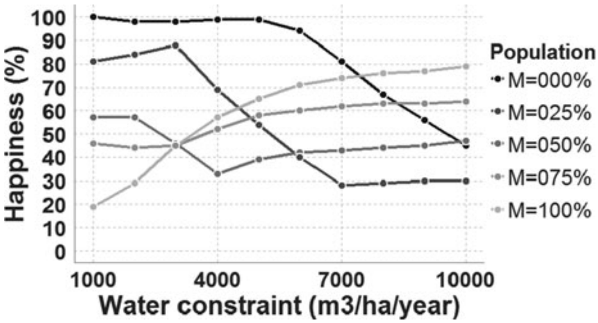


Fig. 35.8 Happiness (%) according to the mind-frames and the water constraint

35.5 Closing Remarks and Future Work

We are interested in understanding the role that values play in the policy-making process and propose the use of agent-based simulation to support value-driven policy design. Our underlying assumption is that policy-makers can achieve a crisper understanding of the consequences of their proposals by making explicit the links between their values and the instruments and expected outcomes they choose. With that assumption in mind, we are convinced that the use of ABS to explore value-driven policy-making will provide a practical tool to support the complex—and usually messy—policy cycle.

In this paper we intend to start substantiating those claims by addressing two specific objectives: first, to outline a conceptual framework that represents the policy-making process as a type of socio-cognitive system, specifically, a type where values are projected onto policy means (afforded actions and their governance mechanisms) and policy ends (metrics and performance indicators that evaluate the effect of agents behaviour on the state of the world). The second objective is to illustrate how agent-based simulation may be used within this framework to model a policy domain; in this case, water use.

Our intention is to further explore policy-making and values in the water policy domain, as it is a rich domain where all the essential components of policy-making are present and, with respect to our research focus, it involves multiple values and multiple stakeholders with different values and the possibility of exploring varied policy schemas [46]. We assume that the socio-hydrologic space can be modelled as a socio-cognitive technical system [31], and therefore the opportunities and challenges in the development of socio-cognitive agents are substantial and non-trivial [15, 31].

In order to further substantiate our claims, we identify the following lines of work:

Reasoning with values and about values: Value aggregation, value comparison, values and norms, values versus goals and preferences.

Developing the notion of policy schema: What are the useful instruments to model? How can means and ends be represented in order to simulate reasoning about them? How are norms related with values?

Modelling the *policy definition* arena: Choosing a reference case. Use of value-based argumentation. Negotiation as conflict resolution? A coherence-based analysis of faction values and actions.

Explore second-order phenomena: Policy adoption generally involves interplay between individual policy subjects and other stakeholders—political factions, special interest groups, and politicians—that react to emerging behaviour and may influence the social space, by sending rhetorical messages to influence other stakeholders, advocating policy changes, introducing incentives, and so on. These phenomena justify bringing some second-order affordances like macro-perception, collective illocutions, and instrument revision into the meta-model.

Addressing empirical questions: Identify domain-specific values, instruments, and indicators. Validation of stakeholder profiles. Development of a "scoring methodology" and value-aggregation models. How to make a simulation model useful.

Explore the usability of the framework: The interplay between the number of afforded actions and values in dispute. Negotiation with ABS support.

Ethical aspects of the use of agent-based models for policy-making: We identify three basic types of issues: (i) Issues involved in the selection and implementation of values for the policy and for the agent's reasoning models (ii) Issues related to the construction of the model itself: assumptions, functionality, and range of application (iii) Issues associated with the use of the agent-based model along the policy cycle. Awareness of the policy-makers about the limitations and scope of system, commitments associated with the use of actual simulations, and so on. Awareness of other stakeholders of the use and validity of the model

We are aware that the design and assessment of public policies entail a multi-disciplinary approach. We trust that our value-driven simulation approach will be attractive to professionals from different fields.

Acknowledgements The first author is supported with the industrial doctoral 2016DI043 grant of the Catalan Secretariat for Universities and Research (AGAUR), sponsored by FCC AQUALIA, IIIA-CSIC, and UAB. This work has been supported by the Catalan-funded AppPhil project (funded by RecerCaixa 2017) and the Spanish-funded CIMBVAL project (funded by the Spanish government, project # TIN2017-89758-R).

References

1. Ahrweiler, P.: Agent-based simulation for science, technology, and innovation policy. Scientometrics **110**(1), 391–415 (2017)
2. Allen, R.G., Pereira, L.S., Raes, D., Smith, M.: Crop evapotranspiration—guidelines for computing crop water requirements. United Nations. http://www.fao.org/3/X0490E/X0490E00.htm, FAO Paper 56 (1998)

3. Atkinson, K., Bench-Capon, T.: Practical reasoning as presumptive argumentation using action based alternating transition systems. Artif. Intell. **171**(10), 855–874 (2007). Argumentation in Artificial Intelligence
4. Banks, J.: Handbook of Simulation. Wiley, Inc., Upper Saddle River (1998)
5. Bannister, F., Connolly, R.: ICT, public values and transformative government: a framework and programme for research. Gov. Inf. Q. **31**(1), 119–128 (2014)
6. Beratan, K.K.: A cognition-based view of decision processes in complex social–ecological systems. Ecol. Soc. **12**(1), 27 (2007)
7. Berger, T.: Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes and policy analysis. Agric. Econ. **25**(2), 245–260 (2001)
8. Berglund, E.Z.: Using agent-based modeling for water resources planning and management. J. Water Resour. Plan. Manag. **141**(11), 04015025 (2015)
9. Boelens, R., Hoogesteger, J., Swyngedouw, E., Vos, J., Wester, P.: Hydrosocial territories: a political ecology perspective. Water Int. **41**(1), 1–14 (2016)
10. Bostrom, A., Fischhoff, B., Morgan, M.G.: Characterizing mental models of hazardous processes: a methodology and an application to radon. J. Soc. Issues **48**(4), 85–100 (1992)
11. Brown, P., Schmidt, J.: Water Ethics: Foundational Readings for Students and Professionals. Island Press, Washington, DC (2010)
12. Camill, P.: Global change. Nat. Educ. Knowl. **3**(10), 49 (2010)
13. Campbell, J.L.: Institutional analysis and the role of ideas in political economy. Theory Soc. **27**(3), 377–409 (1998)
14. Campbell, J.L.: Ideas, politics, and public policy. Annu. Rev. Sociol. **28**(1), 21–38 (2002)
15. Castelfranchi, C.: Minds as social institutions. Phenomenol. Cogn. Sci. **13**(1), 121–143 (2014)
16. Castro, J.E.: Water governance in the twentieth-first century. Ambiente Sociedade **10**(2), 97–118 (2007)
17. DeCaro, D.A., Arnold, C.A.T., Frimpong Boamah, E., Garmestani, A.S.: Understanding and applying principles of social cognition and decision making in adaptive environmental governance. Ecol. Soc. **22**(1), 33 (2017)
18. Dietz, T., Fitzgerald, A., Shwom, R.: Environmental values. Annu. Rev. Environ. Resour. **30**(1), 335–372 (2005)
19. d'Inverno, M., Luck, M., Noriega, P., Rodriguez-Aguilar, J.A., Sierra, C.: Communicating open systems. Artif. Intell. **186**, 38–94 (2012)
20. Galán, J.M., Izquierdo, L.R., Izquierdo, S.S., Santos, J.I., del Olmo, R., López-Paredes, A., Edmonds, B.: Errors and artefacts in agent-based modelling. J. Artif. Soc. Soc. Simul. **12**(1), 1 (2009)
21. Gilbert, N., Ahrweiler, P., Barbrook-Johnson, P., Narasimhan, K.P., Wilkinson, H.: Computational modelling of public policy: reflections on practice. J. Artif. Soc. Soc. Simul. **21**(1), 14 (2018)
22. Gilbert, N., Conte, R., et al.: Artificial Societies: The Computer Simulation of Social Life. Routledge, London/New York (2006)
23. Groenfeldt, D., Schmidt, J.J.: Ethics and water governance. Ecol. Soc. **18**(1), 14 (2013)
24. HM Treasury.: The Magenta Book: Guidance for Evaluation (2011)
25. Jager, W., Edmonds, B.: Policy making and modelling in a complex world. In: Janssen, M., Wimmer, M.A., Deljoo, A. (eds.) Policy Practice and Digital Science: Integrating Complex Systems. Social Simulation and Public Administration in Policy Research, pp. 57–73. Springer International Publishing, Cham (2015)
26. Jones, N.A., et al.: Mental Models: An Interdisciplinary Synthesis of Theory and Methods. Ecol. Soc. **16**(1), 46 (2011)
27. Jones, N.A., et al.: The study of human values in understanding and managing social-ecological systems. Ecol. Soc. **21**(1), 15–24 (2016)
28. Kenter, J.O., et al.: What are shared and social values of ecosystems? Ecol. Econ. **111**(Supplement C), 86–99 (2015)
29. May, P.J.: Policy design and implementation. In: Peters, B., Pierre, J. (eds.) The SAGE Handbook of Public Administration, 2nd edn., pp. 279–291. SAGE Publications, Los Angeles (2012)

30. Moody, R., Gerrits, L.: Introduction to policy-making in the digital age. In: Janssen, M., Wimmer, M.A., Deljoo, A. (eds.) Policy Practice and Digital Science: Integrating Complex Systems. Social Simulation and Public Administration in Policy Research, pp. 205–219. Springer International Publishing, Cham (2015)
31. Noriega, P., Padget, J., Verhagen, H., d'Inverno, M.: Towards a framework for socio-cognitive technical systems. In: Ghose, A., Oren, N., Telang, P., Thangarajah, J. (eds.) Coordination, Organizations, Institutions, and Norms in Agent Systems X. Lecture Notes in Computer Science, vol. 9372, pp. 164–181. Springer, Cham (2015)
32. Pahl-Wostl, C.: Water Governance in the Face of Global Change. Springer International Publishing, Cham (2015)
33. Parker, D.C., Manson, S.M., Janssen, M.A., Hoffmann, M.J., Deadman, P.: Multi-agent systems for the simulation of land-use and land-cover change: a review. Ann. Assoc. Am. Geogr. **93**(2), 314–337 (2003)
34. Parks, L., Guay, R.P.: Personality, values, and motivation. Personal. Individ. Differ. **47**(7), 675–684 (2009)
35. van de Poel, I.: Translating values into design requirements. In: Michelfelder, D.P., McCarthy, N., Goldberg, D.E. (eds.) Philosophy and Engineering: Reflections on Practice, Principles and Process, pp. 253–266. Springer Netherlands, Dordrecht (2013)
36. Polski, M., Ostrom., E.: An institutional framework for policy analysis and design in workshop in political theory and policy analysis (1999), Working Paper No. W98–27
37. Rohan, M.J.: A rose by any name? The values construct. Personal. Soc. Psychol. Rev. **4**(3), 255–277 (2000)
38. Rokeach, M.: The Nature of Human Values. Free Press, New York (1973)
39. Rokeach, M.: Understanding Human Values. Simon and Schuster, New York (2008)
40. Schreinemachers, P., Berger, T.: An agent-based simulation model of human–environment interactions in agricultural systems. Environ. Model. Softw. **26**(7), 845–859 (2011)
41. Schwartz, S.H., Bilsky, W.: Toward a theory of the universal content and structure of values: extensions and cross-cultural replications. J. Pers. Soc. Psychol. **58**(5), 878 (1990)
42. Searle, J.: Rationality in Action. MIT Press, Cambridge (2003)
43. Sengupta, R., Lant, C., Kraft, S., Beaulieu, J., Peterson, W., Loftus, T.: Modeling enrollment in the conservation reserve program by using agents within spatial decision support systems: an example from Southern Illinois. Environ. Plann. B. Plann. Des. **32**(6), 821–834 (2005)
44. Simon, H.A.: Bounded rationality. In: Eatwell, J., Milgate, M., Newman, P. (eds.) Utility and Probability, pp. 15–18. Palgrave Macmillan, London (1990)
45. Stewart, J.: Value Conflict and Policy Change. Palgrave Macmillan, London (2009)
46. Veeneman, W., Dicke, W., Bruijne, M.: From clouds to hailstorms: a policy and administrative science perspective on safeguarding public values in networked infrastructures. Int. J. Public Policy **4**, 414–434 (2009)
47. Vugteveen, P., Lenders, H.J.R., Devilee, J.L.A., Leuven, R.S.E.W., van der Veeren, R.J.H.M., Wiering, M.A., Hendriks, A.J.: Stakeholder value orientations in water management. Soc. Nat. Resour. **23**(9), 805–821 (2010)
48. van der Weide, T.L., Dignum, F., Meyer, J.C., Prakken, H., Vreeswijk, G.A.W.: Practical reasoning using values: giving meaning to values. In: Proceedings of the 6th International Conference on Argumentation in Multi-Agent Systems, ArgMAS'09, pp. 79–93. Springer, Berlin/Heidelberg (2010)
49. Witesman, E.M., Walters, L.C.: Modeling public decision preferences using context-specific value hierarchies. Am. Rev. Public Adm. **45**(1), 86–105 (2015)
50. Zeitoun, M., Warner, J.: Hydro-hegemony—a framework for analysis of trans-boundary water conflicts. Water Policy **8**(5), 435–460 (2006)

Chapter 36
A Philosophical Framework of Shared Worlds and Cultural Significance for Social Simulation

Tom Poljanšek

Abstract In this chapter, I sketch a philosophical framework of shared and diverging worlds and cultural significance. Although the framework proposed is basically a psychologically informed, philosophical approach, it is explicitly aimed at being applicable for agent-based social simulations. The account consists of three parts: (1) a formal ontology of human worlds, (2) an analysis of the pre-semantic significance of the objects of human worlds, and (3) an account of what it means for agents to share a world (or to live in diverging worlds). In this chapter, I will give a brief and concise summary of my account. At the end, I will briefly outline how the proposed framework might be put to use for multiagent social simulation of complex social interaction scenarios involving diverging (cultural) backgrounds.

Keywords Cultural significance · Shared worlds · Social context · Cultural background

36.1 Shared Worlds and Cultural Significance

36.1.1 Motivating the Problem of Diverging Worlds

Within the last few years, the idea that humans live in varying (cultural) *worlds* has gained growing attention among social scientists [1], philosophers [2], and social simulators [3, 4]. However, to date, there is no agreed-upon account of (1) *what (cultural) worlds are* and (2) *how they might or might not be shared* by different agents.

My current work focusses on developing such an account from a philosophical perspective in order to supply social simulators with a valid framework of shared and diverging worlds which might be implemented in the future to improve social simulations of complex social interaction scenarios.

T. Poljanšek (✉)
University of Göttingen, Göttingen, Germany

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_36

36.1.2 A Philosophical Account of Shared and Diverging Worlds

Based on results and considerations from phenomenology, biosemiotics, perception theory, and cognitive science, I developed an account of shared worlds which consists of three parts:

1. An ontology for *human worlds*
2. A theory of the *ecological or cultural significance* of the objects which belongs to human worlds
3. An account of *shared and diverging worlds*

Ad 1) An ontology of human worlds Human worlds are constituted by the objects given to human subjects in ordinary (direct) perception. It is crucial, however, not to identify the phenomenal objects of perception with the mind-independent objects causing the perceptual givenness of phenomenal objects. The objects constituting human worlds are not to be misconstrued as ontological ready-mades awaiting to be grasped by perception. The spatial and temporal boundaries of phenomenal objects are not fixed mind-independently. Rather, the objects constituting human worlds are a positive "achievement" of the subpersonal dispositions and processes underlying human perception [5–7]. Within these subpersonal processes, the sensory stimuli reaching a subject from its nearby physical environment are parsed and amended into *meaningful chunks* (i.e., things, contexts, and events) which are the phenomenal objects of perception [8]. To mark this difference between phenomenal objects given in direct perception and mind-independent physical objects, we have to discriminate two types of ontologies: *ontologies of givenness* (OGs) and *ontologies of existence* (OEs), where the former represent types of objects which constitute the world of a specific subject or a specific type of subjects (e.g., human worlds, animal worlds), while the latter aim at representing types of (physical) objects which exist mind-independently [9, 10]. Accordingly, we must distinguish two ontologically distinct types of objects: *phenomenal objects, which can (only) be given to subjects in (direct) perception* (OOGs), and *physical objects, which exist mind-independently* (OOEs). As I argue elsewhere [10], OOGs neither consist of nor are they reducible to OOEs, although the perceptual givenness of the former depends on the mind-independent existence of the latter. Under normal circumstances, the perceptual givenness of certain OOGs *indicates* the existence of certain OOEs, but OOGs are not comprised of OOEs.

As can further be shown, OOGs given in perception are almost always instances of *object types* the subject in question is already familiar with [10]. Material objects, for example, are immediately perceptually grasped as either (1) *mere material objects* (e.g., a rock), (2) as *artefacts with a certain function or affordance* (e.g., a coffee mug, a flight of stairs, or a door), or (3) as *living things* (e.g., a dog or a person), where the subpersonal classification of OOGs results in the subject having certain expectations about the way an object is normally prone to "behave" with respect to the object type it instantiates (e.g., rocks fall, liquids drip, doors open, and dogs bark). The same goes for the perception of action and event types or for the perception of social contexts and situations: The subject immediately perceives certain types of behavior of other agents *as greetings*, *as fights*, *as a conference*, or *as a breakfast*, where, again, these different types of objects are characterized by specific horizons of expectations comprising the typical developments and "habits" [11] characteristic of the object type instantiated.

A crucial aspect of OOGs is that they are directly and without any conscious reasoning perceived and grasped by human perceivers. A simple example: If you buy some goods in the local supermarket, you do not implicitly or explicitly agree with the cashier that you will now perform a collective action together which consist in you paying and him receiving the money. And nevertheless, without any deferral or complication, you and the cashier realize this collective action. How is that possible? The answer should be obvious: You both immediately and without any consideration grasp *what is going on*, *what to expect of each other*, *what to do next*, and so on. The process of socialization and cultural habitualization has shaped your respective dispositional backgrounds in a sufficiently similar manner, so that you are both disposed to immediately perceive the same OOG (a "payment"), constituted by sufficiently similar (or reciprocal) horizons of expectations [12, 13]. These horizons of expectations also include your respective roles and parts for realizing the interaction type in question [14]. Buying goods is something that you regularly do, taking the money is something the cashier regularly does. That's why both of you are disposed to routinely fulfill what can generally be expected of you, as far as you were raised in a cultural context where *paying for something* is a common action. Thus, in cases such as these (recurring and familiar types of objects, events, contexts, and actions), subjects rely on their tacit cultural backgrounds which prestructure their expectations concerning possible developments of situations and anticipations of their own responsibilities and possibilities for action [10]. Subjects sufficiently familiar with the OOG types in question immediately see OOGs as manifestations of these OOG types: They see a dog *as a dog*, a cashier *as a cashier*, money *as money*, a wedding *as a wedding*, and so on. On the level of experience, the structure of this *perceiving-as* can be explicated through the specific *horizons of expectations* which are immediately (co-)perceived and thus constitutive for the respective OOG types. I call the typical horizon of expectations constitutive of an OOG type its *significance*. OOGs and their significance have thus an anticipatory structure.

Ad 2) The significance of OOGs OOGs are characterized by their *significance* which is best understood as a form of *pre-semantic meaning*. As stated above, the significance of a given OOG (may this be a material object, a social context, or an event) can be explained as the specific *horizon of typical expectations*, *propensities,* or *possibilities* constitutive of the OOG type in question. Thus, for example, part of the significance of a chair as an OOG is that you can sit on it, that it will hold your weight, and so on [15]. Similarly, the cultural significance of a certain gesture or an action type is determined by the way it is normally (or reliably) performed within a certain cultural context, for this reliability enables subjects to internalize respective perceptual dispositions to adequately anticipate what might happen or be done next whenever a specific OOG is perceptually registered.

Generally speaking, the significances of OOGs are ontologically grounded in regularities or structural invariants which hold in a certain physical, mind-independent environment. The dispositional backgrounds of subjects continuously exposed to a specific environment are shaped by such regularities, which thus codetermine the range of OOGs a subject is able to perceive. While a lot of OOGs owe their significance to "higher order structural invariants" in the physical environment (e.g., physical laws) [16], part of the significance of some OOGs – like in the case of gestures, social contexts, and action types – is itself grounded in regularities and reliably performed patterns of human action. In a cultural context like ours, for example, a hand wave can typically be expected to be perceived as a kind of greeting (or farewell) and to be reciprocated by another hand wave of the subject addressed. This is simply the way a hand wave is reliably performed within this cultural context. Thus, the significances of OOGs are not something a subject has to infer consciously, although it must be tacitly familiar with the OOG type in question.

To explain this immediacy of our grasp of OOGs and their significance in more detail, we have to distinguish two types of concepts – *explicit concepts* (ECs) and *tacit concepts* (TCs). ECs and TCs fulfill different functions concerning our epistemic access to the world and reality. While TCs constitute the dispositional background of a subject and thus shape the way it *immediately perceives* the world and the OOGs constituting it, ECs enable subjects to *form explicit propositional attitudes and beliefs* concerning states of affairs in the world (as well as in mind-independent reality). TCs fulfill two functions concerning the way OOGs are immediately given in perception: On the one hand, they dispose subjects to immediately (re)identify different types of objects in perception – call this the *(re)identification function* of TCs. On the other hand, they prestructure the horizon of expectation which is constitutive for the significance of an OOG – call this the *projective function* of TCs. Thus, if a subject has internalized a TC of a certain OOG type, it is able to immediately (re)identify instances of the OOG type perceptively and immediately grasp the respective horizon of expectations that it calls for. This holds for TCs of types of material objects, social contexts, action types, as well as for events. Someone reaches out to you with her right hand, for example, and you directly see *a handshake* and react by immediately reaching out with your hand (where the impulse to do so is realized by the projective function of the TC).

At the same time, to have internalized a TC of a certain OOG type does not necessarily involve having an EC for and the respective beliefs about this object type. To have an EC for a certain object type, you have to be able to form propositional beliefs about the object in question, which is not a necessary condition for perceiving an object as an object of a certain kind by means of your dispositional background. Thus, a dog might have a TC of some kind of handshake while not having an EC for whatever it is doing when it is reaching out with its paw toward a human hand. TCs shape the way we immediately perceive the world, while ECs shape the way we may or may not judge the world and mind-independent states of affairs to be. One important consequence of this idea is that there is always the possibility that the way a person directly perceives a situation and the way the same person judges the situation to be may come apart. A person might, for example, be an avowed anti-racist but still shows immediate aversive reactions toward persons who seem to be foreigners [17].

Ad 3) Shared and diverging worlds Now, the range of OOG types a subject is able to perceive depends on its individual *background of perceptual dispositions* which is, at least in part, the result of its ontogeny and habitualization. Its dispositional background shapes the way the world and its OOGs are immediately perceptually given to a subject. And it also predetermines the typical horizons of expectations which are constitutive for these OOGs. Thus, its dispositional background circumscribes the world of a subject. Now, as far as subjects have similar dispositional backgrounds, they live in the same world, comprised of the same (or sufficiently similar) OOGs bearing the same (or sufficiently similar) significances. Nonetheless, dispositional backgrounds can, at least to a certain degree, differ from subject to subject, depending, for example, on the respective environments they were socialized and habitualized in. Thus, as far as their dispositional backgrounds differ, the worlds subjects live in and the OOGs they perceive do so as well: Subjects might perceive and parse social situations differently, the expectations and tacit rules constitutive for certain action and event types might differ, as might the cultural significance of other OOGs. This, again, can lead to misunderstandings and conflicts emerging from the fact that subjects tend to implicitly assume that the OOGs *they* perceive in a given physical environment are also the OOGs *other subjects* do or should perceive under the same circumstances. On the one hand, humans often tend to reify their direct perception instead of accounting for the possibility of diverging worlds. On the other hand, two or more agents partially share their world as far as their backgrounds overlap. Thus, worldsharing *comes in degrees* [18]. It can be shown, however, that there is a *core of the human world* (or OG), which is shared by all human beings [10]. However, we should not underestimate the degree to which individual worlds of different agents can nevertheless diverge.

36.1.3 Implementing Shared and Diverging Worlds in Social Simulations

Current implementations of cultural dynamics and diverging worlds in social simulations ignore significant aspects of the way human worlds are shared, diverge, and change over time, although important steps into the right direction have already been made within the last few years [1, 3, 4]. However, to adequately model complex social interaction scenarios and cultural dynamics in groups of agents with respect to shared or diverging worlds, the agents in simulations must be endowed with distinct dispositional backgrounds (i.e., sets of TCs), which predetermine the range of OOGs and their respective significances agents are able to perceive. To simplify matters a little bit, the function of TCs might for this purpose be spelled out in a rule-like manner, so that, for example, if an agent perceives a social context *as a funeral* (because she has internalized the respective TC), she changes her behavior (lowers her voice, etc.), while another agent, who has not internalized the respective TC and thus does not "see" *a funeral*, does not do so. By exposure to unfamiliar contexts, event types, or objects, agents might over time internalize new TCs, or the significances associated with existing OOGs might be changed, thus gradually changing their dispositional backgrounds. The second important idea to be implemented in social simulations in order to improve their empirical adequacy is the distinction between OEs and OGs. To simplify matters again, one should distinguish between the mind-independent objects (OOEs) and the OOGs, whose givenness to specific subjects is caused by the former: One and the same hand gesture (thought of as an OOE) might be perceived *as an insult* by one agent or *as an invitation* by another, if their dispositional backgrounds differ, respectively. Thus, we need to model the mind-independent object as well as the diverging OOGs perceived by the agents involved if we want to adequately account for such diversity.

Thus – depending on the specific aim of a social simulation – it becomes necessary to take into account diverging cultural backgrounds as well as their dynamics of experience-based change in order to make social simulations empirically more adequate. I am convinced that the framework, which I develop in more detail elsewhere [10], could be a solid and viable step in this direction and that it is able support social simulators in further improving their most promising approaches.

References

1. Lotzmann, U., Neumann, M.: Simulation for interpretation: a methodology for growing virtual cultures. J. Artif. Soc. Soc. Simul. **20**(3), (2017)
2. Durt, C., Fuchs, T., Tewes, C. (eds.): Embodiment, Enaction, and Culture. Investigating the Constitution of the Shared World. The MIT Press, London (2017)
3. Edmonds, B.: Matching and mismatching social contexts. In: Dignum, V., Dignum, F. (eds.) Perspectives on Culture and Agent-Based Simulations. Integrating Cultures, pp. 149–167. Springer, New York (2014)
4. Pfau, J., Kashima, Y., Sonenberg, L.: Towards agent-based models of cultural dynamics: a case of stereotypes. In: Dignum, V., Dignum, F. (eds.) Perspectives on Culture and Agent-Based Simulations. Integrating Cultures, pp. 129–147. Springer, New York (2014)
5. Hoffmann, D.D.: The interaction of colour and motion. In: Mausfeld, R., Heyer, D. (eds.) Colour Perception. Mind and the Physical World, pp. 361–377. Oxford University Press, New York (2003)
6. Radvansky, G., Zacks, J.: Event Cognition. Oxford University Press, Oxford (2014)
7. Clark, A.: Surfing Uncertainty. Predicition, Action, and the Embodied Mind. Oxford University Press, New York (2016)
8. Smith, B.: Ontologie des Mesokosmos. Soziale Objekte und Umwelten. In: Zeitschrift für philosophische Forschung **52**(4), pp. 521–540 (1998)
9. Brunswik, E.: Wahrnehmung und Gegenstandswelt. Grundlegung einer Psychologie vom Gegenstand her. Deuticke, Leipzig, Wien (1934)

10. Poljanšek, T.: Gemeinsame Welten. Zur Ontologie der menschlichen Wirklichkeit. [In preparation]
11. Husserl, E.: Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie. Biemel, W. (ed.). Husserliana Volume VI. Nijhoff, Den Haag (1976)
12. Bourdieu, P.: Sozialer Sinn. Kritik der theoretischen Vernunft. Suhrkamp, Frankfurt/M (2015)
13. Searle, J.: Wie wir die soziale Welt machen. Die Struktur der menschlichen Zivilisation. Suhrkamp, Berlin (2012)
14. Jackendoff, R.: Language, Consciousness, Culture. Essays on Mental Structure. The MIT Press, Cambridge, London (2007)
15. Gibson, J.J.: The Ecological Approach to Visual Perception. Taylor & Francis, New York (1986)
16. Bermúdez, J.: Thinking Without Words. Oxford University Press, Oxford (2003)
17. Gendler, T.S.: Alief in action (and reaction). Mind Lang. **23**(5), 552–585 (2008)
18. Allen, C.: The geometry of partial understanding. Am. Philos. Q. **50**(3), 249–262 (2013)

Chapter 37
Students of Religion Studying Social Conflict Through Simulation and Modelling: An Exploration

Amrit Bahadur Poudel, Pauline Vos, and F. LeRon Shults

Abstract Researchers at our university use modelling and simulation (M&S) to study religious conflicts, and we wanted to introduce undergraduate students of religion to this research approach. Hence, we started a three-year educational design research project to empirically study ways to introduce these students to M&S as a viable research method in their discipline. The research project will entail several iterations, which aim to have a feasible and effective design of lessons and a better understanding of the learning processes. The first iteration was exploratory and is reported here. For this exploration, we organised a seminar, which was videotaped for post hoc analysis. The seminar started with an introduction of research methods to study violent human behaviour, comparing experiments in which people are exposed to violence in real and virtual worlds. Afterwards, the students explored an agent-based simulation based on Schelling's segregation model. The seminar was concluded as a plenary discussion. After the seminar, follow-up interviews were held with three students. The results suggest that this brief intervention enabled students to gain a good understanding of the way in which M&S can be used to study social conflicts and opened them up to the possibility of adopting the method in their future research. We also found that for an initial understanding of the use of M&S in social research, no knowledge of computational methods is required.

Keywords Students of religion · Research methods · Social conflicts · Social simulations

37.1 Introduction

Research in the social sciences has traditionally been limited to methods such as literature reviews, interviews, observations and surveys. These are also the research methods typically taught to students at universities. However, at our university, we

A. B. Poudel (✉) · P. Vos · F. LeRon Shults
University of Agder, Kristiansand, Norway
e-mail: amrit.poudel@uia.no; pauline.vos@uia.no; leron.shults@uia.no

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_37

have a growing group of researchers studying social phenomena through modelling and simulation (M&S) [1]. It would be beneficial for students to also learn about M&S and not have their courses trail behind innovations. Thus, at our university, we wondered how to introduce students of religion to ways of studying religious conflicts using M&S. We started a three-year educational design research project to study emerging educational approaches that help students reason about how to study violent behaviour in artificial societies (without harming real people). How can we as research practitioners introduce students of religion to new M&S research methods and to inspire some to use these methods in their future research?

Educational design research is a research methodology for designing learning environments in cases where one needs new teaching methods to deal with new content [2, 3]. It enables a scholarly exploration for creating new learning environments, in particular in the area of educational innovation. The intent of the research is to produce design principles of feasible teaching practices and increased understanding on both of those practices and of students' learning. The approach is to start small, evaluate and then iteratively increase the scale of the educational intervention. Each iteration consists of a plan of a sequence of learning activities, which are enacted and formally evaluated. This leads to new reflections, conjectures and improvements of the design. In this chapter, we will present the first iteration of our study.

We hypothesised that an accessible entry into M&S research methods for students of religion would not be through the learning of coding and a crash course on the creation of M&S [4]. Instead, students of religion could better think of M&S as virtual worlds of social phenomena, which can imitate or reproduce real-world processes. We hypothesised that they could imagine that virtual worlds can support social researchers to run social experiments or to see what future scenarios could occur. We selected a digital simulation based on authentic research of human behaviour (described below) to serve as an example of M&S research. We planned it to potentially engage students in critical thinking, imagining alternative actions and problem-solving skills.

In our research on learning, we perceive students not only as actors in a learning context (as learners and as peers within a university course), as members of the university community, but also as social beings with extra-institutional experiences in the society. Within the narrow learning context, we use probing questions and a simulation to mediate student thinking, participation and communication. We perceive M&S as cultural tools that carry knowledge about social phenomena and hence capable of incorporating a social history into any learning activity in which religion students engage. We study their learning not simply as an interplay between minds and simulations; rather, we frame students learning in light of three sociocultural contexts: (1) the course, (2) the world of academic researchers using M&S in social research and (3) wider society in which there is a need to understand and limit conflicts. This theoretical frame of learning is inspired by Engeström's Cultural–Historical Activity Theory (CHAT) [5], which was developed to analyse the relationship between what people think, what they do, how they use tools and how different contexts interact in this. Our questions were: What are the strong and weak points in our educational design? To what extent can the students understand how others (social researchers at their university) use M&S in their research? Would they using this approach in their future research?

37.2 Materials and Participants

In this first exploration of our design-based study, we created a learning context that consisted of a three-hour seminar that students could attend on a voluntary basis. It was an ‘extra’ within the compulsory bachelor’s course ‘Religious radicalisation, extremism and violence’ taught by a group of lecturers from the department. The seminar was conducted by the second author. All students had been asked to bring a laptop. Those who hadn’t brought one were paired to a peer with one. All activities in this seminar were filmed by the first author. All three authors participated in data analysis.

The seminar consisted of three sections. The first section was a lecture, in which different research methods in the social sciences were reviewed. This was followed by a discussion of the problem of carrying out live laboratory experiments with people. We presented the Zimbardo prison experiment carried out at Stanford University, designed to investigate whether random volunteers had dispositions associated with abusive behaviour [6]. This prison *simulation* has become famous, because it had to be stopped due to the excessive abuse that arose. It serves as an example of how live experiments can have serious ethical issues. Alternatively, in artificial societies, where individuals are not actual people but virtual entities, we can experiment with scenarios in ways that would be unethical in real societies with real people (Fig. 37.1).

In the second part of the seminar, we gave students hands-on experience with a simulation. For this, we selected a simulation that was freely available online, had a limited number of parameters (sliders) and connected to a theme that

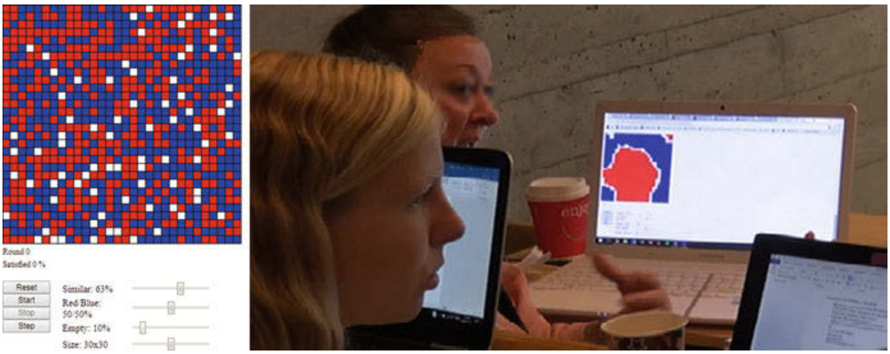


Fig. 37.1 Simulation based on Schelling’s model of segregation (left) and classroom scene (right)

was likely to be of interest of the students: the social inclusion and exclusion of people. The simulation was based on a model of segregation created by American economist Thomas Schelling (http://nifty.stanford.edu/2014/mccown-schelling-model-segregation/). This simulation engages a user in understanding the processes of the segregation of two recognisable groups. The bottom line of this model is that segregation can emerge even in relatively tolerant populations. It shows connections between micro-motives and macro-behaviour and reveals *hidden causal architectures* in social systems. The third part of the seminar consisted of a discussion guided by probing questions such as: Why do researchers use M&S? What questions could be answered by creating a virtual Norway? What are limitations of M&S? Are you interested in doing research using M&S?

In the weeks after the seminar, the first author conducted follow-up interviews with three students, who volunteered to be interviewed. They were asked to evaluate the seminar and describe what they recalled from it.

The video recordings of participants' interaction in the seminar (with the simulation and with each other) and the videos of the follow-up interviews were analysed in light of the CHAT framework with the goal of answering the research questions.

37.3 Results

Interaction of the students during the seminar Students fully engaged in the hands-on activity and discovered how small individual bias can lead to large collective segregation. They interacted with the simulation by using the sliders and click buttons. We observed wonder, excitement and sorrow on their faces. They played with the parameters to explore future scenarios, and without exception, they tried to create positive, unsegregated outcomes. This shows that within a short time, the simulation had become a tool, which students manipulated to reach a certain goal, and that it appealed to their social feelings. Thus, the simulation created a connection between the students, their goals and social life beyond the university.

In the third part of the session, the teacher posed probing questions. To the question on what could be studied using M&S, students came up with several examples. It could be applied to forecasting elections, to predicting unemployment rates, 'to understand criminality by understanding people's behaviour' or to discover 'what would happen if radical religious groups came into power'. These varied answers indicate that the students were able to connect M&S to doing social research. Additionally, two students expressed interest in using M&S in future research.

Follow-up interview with students In the follow-up interview, one student said: 'It is easy to find answers to hypothetical questions in social research using social simulation..... we may not be able to afford experiments like Zimbardo, which has a high cost value as well it as affects peoples' personal lives. Instead, if you run computer simulations, it is less harmful and more cost effective.' Another student reflected: 'It made me easier to understand when LeRon talked about computer simulations later on in the course.' These answers indicate that the students were able to connect the simulations to researchers who use M&S. Moreover, when a student uses the pronoun 'we', this suggests he considers himself a potential researcher.

37.4 Conclusions, Discussion and Recommendations

We observed that the students of religion showed an understanding of the opportunities M&S offer to understand hidden architectures of social phenomena. They were very aware that live experiments with people have disadvantages, and they could imagine questions that only M&S can answer. During the seminar, there was increased understanding that M&S can be a tool in research on religious conflicts. Some students showed a keen interest in M&S for their future research. We conclude that the M&S approach to studying social phenomena offers a pedagogical innovation for dealing with topics such as religious radicalisation, extremism and violence. Of course, it remains speculative whether this first encounter with M&S will motivate them to take up a course in programming. We will further explore this in the next iterations of our research.

References

1. Shults, F.L., Lane, J.E., Wildman, W.J., Diallo, S., Lynch, C.J., Gore, R.: Modelling terror management theory; computer simulations of the impact of mortality salience on religiosity. Religion, Brain & Behav. **8**, 77–100 (2018)
2. Barab, S., Squire, K.: Design-based research: putting a stake in the ground. J. Learn. Sci. **13**, 1–14 (2004)
3. McKenney, S., Reeves, T.C.: Educational design research. In: Handbook of Research on Educational Communications and Technology, pp. 131–140. Springer, New York (2014)
4. Padilla, J.J., Lynch, C.J., Kavak, H., Evett, S., Nelson, D., Carson, C., del Villar, J.: Storytelling and simulation creation. In: 2017 Winter Simulation Conference (WSC), pp. 4288–4299. IEEE (2017)
5. Engeström, Y.: Studies in Expansive Learning: Learning What Is Not Yet There. Cambridge University Press, Cambridge (2016)
6. Carnahan, T., McFarland, S.: Revisiting the Stanford Prison Experiment: could participant self-selection have led to the cruelty? Personal. Soc. Psychol. Bull. **33**, 603–614 (2007)

Chapter 38
Using Cognitive Work Analysis to Inform Agent-Based Modelling of Automated Driving

Gemma J. M. Read, Paul M. Salmon, and Jason Thompson

Abstract Cognitive work analysis is a systems analysis framework from the field of human factors. The framework provides information about the constraints of agent behaviour that can potentially be useful for the development of agent-based models. This chapter discusses a current programme of research which aims to integrate cognitive work analysis with agent-based modelling to identify and manage emergent risks to safety associated with the introduction of automated vehicles into the road transport system. A proposed approach to the integration of the methods is described.

Keywords Agent-based modelling · Cognitive work analysis · Complex systems · Automated vehicles

38.1 Introduction

The aim of this chapter is to propose the use of a framework from the field of human factors to inform the development of agent-based models (ABMs). Specifically, we discuss a current programme of research which aims to integrate cognitive work analysis (CWA) with ABM to identify and manage emergent risks associated with the introduction of automated vehicles (AVs). Road transport systems continue to move towards fully automated driving. However, the shift towards full automation

G. J. M. Read (✉) · P. M. Salmon
Centre for Human Factors and Sociotechnical Systems, University of the Sunshine Coast, Maroochydore, Australia
e-mail: gread@usc.edu.au

J. Thompson
Transport, Health and Urban Design (THUD) Research Hub, Faculty of Architecture, Building and Planning, Melbourne School of Design, The University of Melbourne, Melbourne, VIC, Australia

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_38

is occurring within already complex and poorly understood road systems that kill people on a scale comparable with other major public health issues such as cancer and cardiovascular and respiratory diseases.

Whilst it is expected that automation will create safety and efficiency benefits, it also poses risks [1]. Importantly, little is known about how other road users (e.g. human drivers, cyclists, and pedestrians) will respond and adapt to sharing the road with AVs. Indeed, the introduction of advanced automation in road transport brings new challenges, different to those faced in other safety critical domains, such as aviation, which have already transitioned. For example, in road transport, automated entities will interact directly with members of the public who do not have specific training, nor close supervision, to manage their reactions and behaviours. It is questionable the extent to which AV designers can take account of the openness and complexity of the road system, making this an important area for exploration.

38.1.1 Cognitive Work Analysis

CWA provides a framework for analysing complex systems [2]. It was developed at the RISØ National Laboratory in Denmark within a wider research programme aimed at supporting the development of safe electricity from nuclear power. The researchers found that even where reactors could be designed to exhibit close to perfect technological reliability, accidents still occurred. The studies at RISØ culminated in the emergence of the CWA framework of tools to assist in the design of adaptive systems that enabled the worker to 'finish the design' [2].

The CWA framework is unique in its formative, constraint-based approach that models the possibilities for behaviour within complex systems, rather than describing actual behaviour (i.e. how work is done) or prescribing normative behaviour (i.e. how work should be done) [2]. It has been widely used to analyse complex systems including in nuclear power generation, military command and control, air traffic control, health care, road transport and rail transport. CWA uses five phases of analysis to analyse the constraints on agent behaviour (both human and non-human) within a system:

1. Work domain analysis which uses the abstraction hierarchy or abstraction–decomposition space to identify and describe the functional purpose and structure of the work domain.
2. Control task analysis which uses the decision ladder and contextual activity template to identify and describe the activities and tasks performed in the system.
3. Strategies analysis which applies information flow diagrams to identify the strategies that can be employed to perform the activities and tasks.
4. Social organisation and cooperation analysis which uses annotated versions of the above-mentioned representations to identify how tasks and activities are distributed across agents within the system.

5. Worker competencies analysis which uses the skills, rules and knowledge taxonomy to identify the cognitive skills and processes employed during task performance.

CWA enables the analyst to identify the constraints or boundaries on the possibilities for action within a given system. These can be hard constraints (e.g. those imposed by the laws of physics) or soft constraints (e.g. road rules, social norms). Although the constraints identified by CWA are situation dependent and dynamic [2], CWA is not able to formally represent this dynamism. Further, analyses to identify the full range of behavioural strategies available in a given situation using CWA tools can quickly become large and unwieldly. Instead, the constraints analysed with CWA could be used to define aspects of dynamic models (such as rules for interaction), enabling emergent behaviours to be identified. Also, being a qualitative approach, CWA could be enhanced through integration with quantitative modelling approaches.

38.1.2 Agent-Based Modelling

By contrast to CWA, ABM uses a bottom-up approach to modelling complex systems. In ABM, rules are applied to direct the behaviour of individual agents who then interact, enabling the identification of emergent patterns and phenomena. ABM demonstrates how agents self-organise based on relatively simple rules of interaction, providing unique insight into how interactions can be changed or managed to avoid emergent patterns that can lead to road crashes. The approach is highly suited to forecasting patterns of behaviour in future road systems where vehicles operating at different levels of automation will interact with one another and with vulnerable road users (e.g. pedestrians and cyclists). ABM has previously been applied to transportation systems [3–5]. Applications of ABM to the area of AVs are emerging, yet tend to focus on a homogenous fleet. For example, ABM has been used to develop and test algorithms for efficient and safe interactions between AVs [6]. There have been calls for improving the theoretical underpinning for understanding interactions between agents during the transition period and for improving the integration of human factors within models [7].

38.2 Proposed Approach to Method Integration

Both CWA and ABM seek to understand adaptation in complex systems and how phenomena emerge from interactions between system agents (both human and non-human). It is expected that combining the two approaches will provide improvement over existing methods as it will describe agent interactions from a constraint-based systems perspective (enhancing the capability of ABM approaches) and

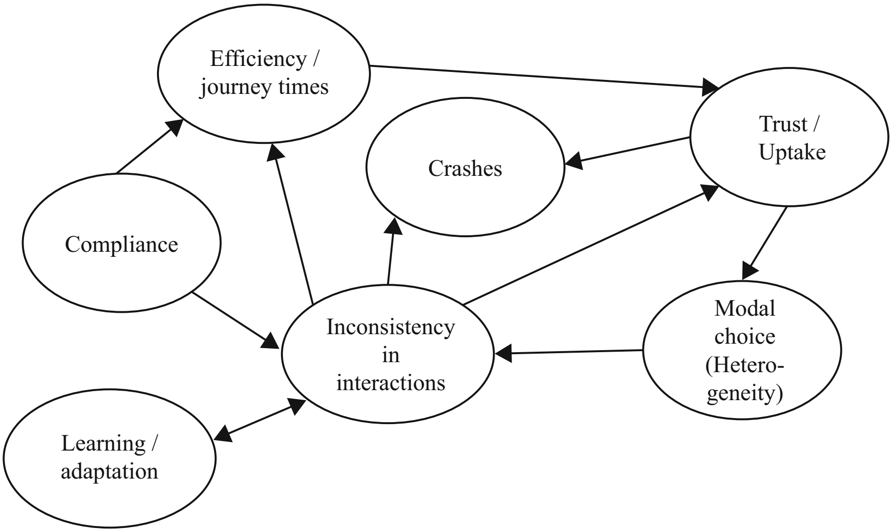


Fig. 38.1 Conceptual model

dynamically modelling interactions to identify emergent risks with the introduction of new technologies (enhancing the capability of CWA). Indeed, previous work has taken steps to use CWA to define agent-based simulations of agent activities [8]. Using the two approaches also provides an opportunity to validate the output of each (i.e. are the emergent risks identified by ABM also identified by CWA).

Our approach to integrating ABM and CWA is both high level and detailed. Initially, we have identified a high-level conceptual model to underpin the modeling (Fig. 38.1). The variables included in the model were derived from a work domain analysis of the road transport system [9], adapted for AV introduction into the road system. The conceptual model describes key relationships between model choice, which either increases or decreases heterogeneity in the traffic fleet, which impacts on the extent to which there are inconsistencies in interactions between road users. For example, where pedestrians encounter very few AVs, their expectations regarding interactions will only be potentially violated in small number of interactions. As there is more uptake of AVs, the number of inconsistencies in interactions will increase; however, this will be mediated by learning through experience, which will change expectations and reduce inconsistent interactions. Compliance will also impact on the inconsistency of interactions. AVs will comply with road rules which should assist other road users to predict their behaviour. However, this may lead other road users to take advantage of AVs, such as pedestrians crossing in front of traffic, knowing that AVs will stop for them. This has implications for efficiency of AVs and of the road network in general, which could reduce trust and their uptake by consumers. Finally, some types of inconsistent interactions could lead to crashes. These incidents will also reduce trust and the uptake of AVs.

The conceptual model will be used to develop a high-level model to explore the broad mechanisms behind crashes and the kinds of scenarios/combinations of scenarios that lead to crashes. Then, more detailed models will be created to understand key areas of risk. For the more detailed models, we are developing subsystem CWA models of various road environments such as signalised intersections, roundabouts, and motorways. The CWA outputs will be used to specify the agents to be included in the model, their attributes and the environment within which agents will interact. Importantly, the analysis will identify the rules determining interactions between agents and between agents and the environment. For example:

- Work domain analysis (Phase 1 of CWA) identifies the environmental constraints on agent behaviour. This can assist in defining environmental elements of models, such as the use of lane markings as a constraint on agents' movement through space.
- Strategy analysis (Phase 3 of CWA) identifies the different behavioural strategies that agents can use to undertake tasks. This can assist to define potential agent behaviours such as the various ways in which cyclists can proceed through intersections (in traffic flow behind AV, alongside AV, on footpath, via hook turn).
- Worker competency analysis (Phase 5 of CWA) identifies the cognitive capabilities (skill, rule and knowledge based) that agents require to undertake tasks. This can assist to define agent rules, such as 'IF vehicle approaches from right THEN give way at intersection'.

A key challenge will be to avoid overspecifying the detailed models, such that it becomes simply a dynamic simulation of what was found through the CWA. The careful selection of key constraints and attributes will be important. Further, CWA currently operates on the basis of the system at a point in time. Whilst it allows for dynamic shifts in behaviour, it does not necessarily attempt to predict or model them. Within the ABMs, learning algorithms, based on human factors theories such as the perceptual cycle model [10], will be incorporated to understand how interactions might change over time.

38.3 Conclusions

The introduction of AVs into the existing road transport system poses potential safety challenges. We propose the integration of ABM and CWA to understand potential system phenomena and safety risks to support the safe introduction of AVs.

References

1. Stanton, N.A., Marsden, P.: From fly-by-wire to drive-by-wire: safety implications of automation in vehicles. Saf. Sci. **24**(1), 35–49 (1996)
2. Vicente, K.: Cognitive Work Analysis. Lawrence Erlbaum Associates, Mahwah (1999)
3. Correa, J.C.: The behavioral interaction of road users in traffic: an example of the potential of intelligent agent-based simulations in psychology. Rev. Latinoam. Psicol. **41**, 201–208 (2016)
4. Kim, S., et al.: Impact of road environment on drivers' behaviors in dilemma zone: application of agent-based simulation. Accid. Anal. Prev. **96**(1), 329–340 (2016)
5. Thompson, J., Newnam, S., Stevenson, M.: A model for exploring the relationship between payment structures, fatigue, crash risk, and regulatory response in a heavy-vehicle transport system. Transp. Res. A. **82**, 204–215 (2015)
6. Mladenović, M.N. & Abbas, M.M: Self-organizing control framework for driverless vehicles. Proceedings of ITSC **2013**, The Hague, Netherlands (2013)
7. van Lint, H., et al.: Getting the human factor into traffic flow models – a new open-source design to simulate next-generation traffic operations. Transp. Res. Rec. **2561**, 25–33 (2016)
8. Sanderson, P., et al.: Use of cognitive work analysis across the system life cycle: from requirements to decommissioning. HFES Ann. Meeting. **43**, 318–322 (1999)
9. Salmon, P.M., Read, G.J.M.: Many model thinking in systems ergonomics: a case study in road safety. Ergonomics. **62**, 612 (2019)
10. Neisser, U.: Cognition and Reality: Principles and Implications of Cognitive Psychology. Freeman, San Francisco (1976)

Chapter 39
Modelling the “Captain’s Nose”: Exploring the Shift Towards Autonomous Fishing with Social Simulation

Jorge Santos, Melania Borit, and Loïs Vanhée

Abstract With the rapid advances in the development of autonomous vehicles, the fishing industry might be faced in the near future with a shift to (partially or fully) autonomous fishing vessels or operations. Large-scale utilization of autonomous fishing practices will lead to a reorganization of society in many respects, with critical challenges related to conservation, economy, governance, and ethics. This study focuses on some relevant topics to reflect on with social simulation approaches. These topics will have to be revisited when exploring the possible challenges and opportunities of moving from traditional to autonomous fishing activities.

Keywords Fishing industry · Human-agent interaction · Interdisciplinary approaches · Maritime robotics · Multi-agent systems · Social modelling

39.1 Introduction

Even though partially or fully autonomous vehicles are present-day technology in several land based or marine transport modes (e.g. [3, 5, 17, 35]) and are also explored as a possibility for maritime operations (e.g. [12, 27, 28, 31]), for various reasons, notably cultural [26], the domain of fishing has never been explored until now from this perspective. Fishing can be defined as any activity that involves the catching, taking, or harvesting of fish [13], and it basically involves the existence of a prey (i.e. the fish) and of an operator (i.e. the fisher) who uses some form of fishing gear. At sea, fishing operations are typically supported by specialized vessels. Fishing is the core activity of fisheries, which is a concept analysed lately from the

J. Santos
UiT The Arctic University of Norway (University of Tromsø), Tromsø, Norway

M. Borit (✉)
Norwegian College of Fishery Science, UiT The Arctic University of Norway, Tromsø, Norway
e-mail: melania.borit@uit.no

L. Vanhée
GREYC University of Caen, Caen, France

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_39

perspective of socio-ecological complex adaptive systems [21]. Regardless of the challenges of sustainable fisheries management [7], fish is one of the most traded food commodities in the world [10]. By, among others, increasing human safety, providing access to remote or difficult areas (e.g. rough weather spots, deep sea), or decreasing costs associated with ship size, moving towards autonomous fisheries might be seen as beneficial, especially in fisheries targeting valuable species where envisaged profits might just push new technology through the cultural barriers mentioned by [26]. The future of such autonomous fishing (AF) operations can be easily imagined from a technical point of view, with multi-agent approaches called in to address the difficulties rising from transforming a classic fishing vessel into a partially or fully AF vessel, operating alone or together with supporting marine autonomous vehicles (underwater, surface, or aerial vehicles) in an AF operation system (AFOS) [37]. However, many of the challenges of the transition from traditional fishing to AF stem from the dynamics of social interaction and the fine mechanisms of social processes and, thus, can be explored using social simulation (SocSim) approaches. Without focusing on a specific type of fisheries, we analyse here several classic SocSim topics that we consider immediately relevant.

39.2 Historical Background and a Look into the Future

The history of fisheries in the twentieth century shows that technology adoption in this domain was quick, but that it took different paths in different ecosystems. For some of the new technologies, the fishing industry was not only an early taker, but also a strong pressure group for their consolidation (e.g. differential GPS). All the technology build-up (e.g. gradual adoption of radars, echosounders, sonars, GPS), combined with deficient regulations or policies, had the downside of the large increase in industrial fishing power [30] with a corresponding dwindling of many important fish stocks. The increased regulation of the last 30 years has led to the decommission of many old vessels (and fishers), which have often been replaced by fewer, but larger vessels. This upscaling is also an innovation, with comprehensive changes in labour relationships, financial organization, processing, and marketing. It became soon clearer that technology was replacing tradition, skill, and human intuition (the famous "captain's nose" [6]): in modern fishing it is not the skipper's skill and gut feeling, but more the size of the boat and experience of the crew that contribute most to fisher's success [25]. With the advent of autonomous vehicles in many other related areas, the fisheries system might be facing a revolution with far reaching environmental and social consequences. Despite social and technological developments, fisheries are still, and will be, very diverse activities owing to the different ecological and geographical settings. The new type of AF will be increasingly embedded with marine aquaculture, but also with all other social and economic activities happening at sea and on land, with technology augmenting the capacity of humans to deliver niche products. The use of this new technology might lead to setting the ground for a renaissance of clusters along the coasts, a vision that is more related to peer-to-peer economy and a decentralized, regional territoriality [14] (though the opposite might also happen, with stocks being exploited by supranational shing conglomerates).

39.3 Topics to Be Revisited

Technological shifts and acceptance of new technology Innovation and acceptance of new technology depend on the cultural context (e.g. kinship), the markets and the value chains under which fishers operate, the age of the operators and their long-term commitment to the fishery, as well as the regulatory environment [1, 22, 33]. For thousands of years, fisheries were dominated by seasonal activities where traditional knowledge and kin relations in coastal communities played a central role. Early in the industrialization of fisheries, by the 1930s, fishing operators still possessed a general all round set of skills and knowledge, which were passed on within families and communities, but they were being embedded in a new technological and institutional complexity. By the year 2000, in the industrial countries, and increasingly around the world, the main activity of a fisher is to run a business, control finances and investments, and understand fishing legislation [18]. The fishing knowledge itself is more embedded in technology systems, and organizations, and the important roles of kin have been replaced by technical education and often imported labour. Technological innovation and adoption goes hand in hand with the development of new processing technology and markets, as well as institutional innovation (e.g. cooperatives, mutuals, or other forms of future governance). There is clearly a pressing need to understand the cultural and social relations that will hinder, allow, or favour the adoption of new technology and (full) autonomy in the different fisheries and ecosystem settings. SocSim offers a strong approach for evaluating the consequences of this technological transition on societies [23], and previous research in areas such as social norms, social influence, values/culture, trust, social psychology, reputation, and social networks will play an important role here (e.g. [9, 24, 34]). Moreover, serious gaming and participative modelling could facilitate this transition [4].

Resource exploitation In future fisheries, production might be more demand-driven (probably for top-quality live fish) and autonomous agents will have to take decisions as to how and when to search for fish (i.e., prey). Once a patch of prey is encountered, decisions will have to be made as to how long to stay there before moving on. Predator-prey models have a long history in ecology and have been amply dealt with in individual-based modelling [16]. This includes dealing with the typical patch decision models that play an important role in the optimal foraging theory [8]. Like any other predator, automated agents will have to compound the benefits and costs, including opportunity costs, of staying or leaving a patch, as well as of cooperating with other agents in a competitive market economy. Differently from traditional resource exploitation models, the agents must deal, or learn to deal, with different types of prey, regulations, variable order flows, information sharing, and new governance structures. This socio-techno-ecological complex adaptive system of fisheries constitutes a relevant opportunity for establishing new agent-based models, which can be applied for predicting the relevance of combinations of agents, policies, and institutional structures.

Governance AFOSs will include multiple sensors and nearly perfect knowledge of their local environment (e.g. schools of prey). It will be important to match production capacity of a legion of competing AFOSs, spread over a large area, to the stock size and productivity of the whole prey stock. Owing to the efficiency of autonomous agents as predators, the information and (backward) control flows must happen in real time, to ensure environmental sustainability. This raises issues about the veracity and asymmetry of the information shared in the network, as well as trust and maintenance of trade secrets (commercial advantages) within competing firms. Should networks be controlled by a principal (the government) or multiple principals (agencies) or delegated to trusted agents or to peer-to-peer mechanisms? Application of design thinking to the study and development of optimal governance structures is a relatively recent field, and SocSim is a promising approach to experiment and compare these different control regimes [19].

Value chains/network AF will, in principle, be less constrained by time or weather limitations. This could be suitable for establishment of small supplier firms and for the production and trade of live fish and top quality products. This translates into better welfare for the fish and a superior market for the supplier. In a globalized world, there is less competition between individual actors than between value chains. These chains tend to become shorter and demand-driven. The configuration of the chains, type of product, and markets will dictate whether these chains will be dynamic or stable, and this can be explored by SocSim [2]. It is relevant to analyse in which conditions vertical integration, marketing and private certification schemes, strategic alliances, or horizontal arrangements, such as cooperatives, peer-to-peer networking, or other hybrid models, will be most suitable. Previous SocSim research in areas such as negotiations, coordination, cooperation, reputation, social networks, or learning will be highly relevant here (e.g. [11, 32, 36]).

Values/morals/ethics Stakeholders and consumers are increasingly conscious of other values than just business or social economics (such as social responsibility, animal welfare, or valuing nature for its own sake). Simulations on AFOSs supply chains should go further and highlight new dividing issues, some of which might have unexpected features. How long should the fish be kept alive after capture and before final consumption to ensure freshness and value, but avoid prolonged stress? Should AF only target the highly valued species for the rich consumer markets? What balance of opportunities do we want between large-scale industrial and small-scale community fisheries?

Modelling Focusing SocSim efforts on this complex physical, social, and technological system in transition could contribute to learning how to make better models. From this perspective, one key challenge would be eliciting the necessary (and many times implicit) information from the fishers in order to (1) programme the AFOS and its interaction with humans and (2) programme the SocSim models needed to explore the topics mentioned above. Participatory modelling (e.g. [20]) and efforts towards devising methodologies for using qualitative data for informing agent rules (e.g. [15]) would play a crucial role here.

39.4 Conclusion

The transition to AF both depends on and is consequential on society as a whole. Thus, designing social simulations could be a useful tool for a variety of purposes, from making decisions about, for example, policies or investments, to exploring the consequences on society and organizations of the evolution of fishing techniques and fishing activities towards autonomy. Such endeavours might need better social theories and data [29]. Nevertheless, they could also contribute to formulating new theories and uncovering undesirable outcomes and promising futures.

References

1. Acheson, J.M., Reidman, R.: Technical innovation in the New England fin-fishing industry: an examination of the Downs and Mohr hypothesis. Am. Ethnol. **9**(3), 538–558 (1982)
2. Akanle, O.M., Zhang, D.Z.: Agent-based model for optimising supply-chain configurations. Int. J. Prod. Econ. **115**(2), 444–460 (2008)
3. Bahnes, N., Kechar, B., Haffaf, H.: Cooperation between intelligent autonomous vehicles to enhance container terminal operations. J. Innov. Digit. Ecosyst. **3**(1), 22–29 (2016). https://doi.org/10.1016/j.jides.2016.05.002
4. Barreteau, O., Bousquet, F., Attonaty, J.M.: Role-playing games for opening the black box of multi-agent systems: method and lessons of its application to Senegal River Valley irrigated systems. J. Artif. Soc. Soc. Simul. **4**(2), 5 (2001)
5. Batalden, B.M., Leikanger, P., Wide, P.: Towards autonomous maritime operations. In: 2017 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), pp. 1–6. IEEE (2017). https://doi.org/10.1109/CIVEMSA.2017.7995339
6. von Brandt, A., Others: Fish catching methods of the world. Fishing News Books, Farnham (1984)
7. Costello, C., Ovando, D., Clavelle, T., Strauss, C.K., Hilborn, R., Melnychuk, M.C., Branch, T.A., Gaines, S.D., Szuwalski, C.S., Cabral, R.B., Others: Global fishery prospects under contrasting management regimes. Proc. Natl. Acad. Sci. **113**(18), 5125–5129 (2016)
8. DeAngelis, D.L., Mooij, W.M.: Individual-based modeling of ecological and evolutionary processes. Annu. Rev. Ecol. Evol. Syst. **36**, 147–168 (2005)
9. Dignum, F., Prada, R., Hofstede, G.J.: From autistic to social agents. In: Proceedings of the 2014 International Conference on Autonomous Agents and Multi-Agent Systems, pp. 1161–1164. International Foundation for Autonomous Agents and Multiagent Systems (2014)
10. Duddu, P.: The ten most traded food and beverage commodities (2014). http://www.foodprocessing-technology.com/features/featurethe-10-most-traded-food-and-beverage-commodities-4181217/

11. Edmonds, B., Meyer, R.: Simulating Social Complexity: A Handbook. Springer, Cham (2017)
12. Edwards, D.B., Bean, T.A., Odell, D.L., Anderson, M.J.: A leader-follower algorithm for multiple AUV formations. In: 2004 IEEE/OES Autonomous Underwater Vehicles, pp. 40–46. IEEE (2004). https://doi.org/10.1109/AUV.2004.1431191
13. FAO: FAO Fisheries Glossary (2017). http://www.fao.org/faoterm/viewentry/en/?entryId=98375
14. Gál, Z., Lux, G.: ET2050 Territorial Scenarios and Visions for Europe. Technical report, Final Report 30 (2014)
15. Ghorbani, A., Dijkema, G., Schrauwen, N.: Structuring qualitative data for agent-based modelling. J. Artif. Soc. Soc. Simul. **18**(1), 2 (2015)
16. Grimm, V., Wyszomirski, T., Aikman, D., Uchmański, J.: Individual-based modelling and ecological theory: synthesis of a workshop. Ecol. Model. **115**(2–3), 275–282 (1999)
17. Ioannou, P.: Automated Highway Systems. Springer Science & Business Media, New York (2013)
18. Johnsen, J.P., Holm, P., Sinclair, P., Bavington, D.: The cyborgization of the fisheries. On attempts to make fisheries management possible. MAST. **7**(2), 9–34 (2009)
19. Koliba, C., Zia, A.: Governance Informatics: Using Computer Simulation Models to Deepen Situational Awareness and Governance Design Considerations. MIT Press, Cambridge (2013)
20. Le Page, C., Perrotton, A.: KILT: a modelling approach based on participatory agent-based simulation of stylized socio-ecosystems to stimulate social learning with local stakeholders. In: International Conference on Autonomous Agents and Multiagent Systems, pp. 31–44. Springer (2017)
21. Levin, S., Xepapadeas, T., Crépin, A.S., Norberg, J., De Zeeuw, A., Folke, C., Hughes, T., Arrow, K., Barrett, S., Daily, G., Others: Social-ecological systems as complex adaptive systems: modeling and policy implications. Environ. Dev. Econ. **18**(2), 111–132 (2013)
22. Mazuki, R., Man, N., Omar, S.Z., Bolong, J., D'Silva, J.L., Azril, H., Shaffril, M.: Technology adoption among fishermen in Malaysia. J. Am. Sci. **8**(12), 1–4 (2012)
23. Mccarthy, J.G., Sabbadini, T., Sachs, S.R.: Multi-agent model of technological shifts. Work **24**, 112–127 (2008)
24. Nooteboom, B.: Agent-based simulation of trust. In: Lyon, F., et al. (eds.) Handbook of Research Methods on Trust, pp. 65–74. Edward Elgar Publishing, Cheltenham (2015)
25. Pálsson, G., Durrenberger, P.: To dream of fish: the causes of Icelandic skippers' fishing success. J. Anthropol. Res. **38**(2), 227–242 (1982)
26. Pieters, K.: The Near Future of Unmanned Vessels. A Complexity-Informed Perspective (2017). https://www.researchgate.net/publication/316364255_The_Near_Future_of_Unmanned_Vessels_A_Complexity-Informed_Perspective
27. Poedjono, B., Pai, S., Maus, S., Manoj, C., Paynter, R., Others: Using autonomous marine vehicles to enable accurate wellbore placement in the Arctic. In: OTC Arctic Technology Conference. Offshore Technology Conference (2015)
28. Price, R.R., Hall, S.G.: Design, development, and testing of an autonomous boat to reduce predatory birds on aquaculture ponds. Biol. Eng. Trans. **5**(2), 61–70 (2012). https://doi.org/10.13031/2013.41399
29. Raser, J.R.: Simulation and Society: An Exploration of Scientific Gaming. Allyn and Bacon, Boston (1969)
30. Robins, C.M., Wang, Y.G., Die, D.: The impact of global positioning systems and plotters on fishing power in the northern prawn fishery, Australia. Can. J. Fish. Aquat. Sci. **55**(7), 1645–1651 (1998)
31. Salamon, A., Housten, D., Drewes, P.: Increasing situational awareness through the use of UXV teams while reducing operator workload. J. Field Rob. **25**(9), 598–614 (2008)
32. Squazzoni, F.: Agent-Based Computational Sociology. Wiley, Hoboken (2012)
33. Squires, D., Vestergaard, N.: Technical change in fisheries. Mar. Policy **42**, 286–292 (2013)
34. Stummer, C., Kiesling, E., Günther, M., Vetschera, R.: Innovation diffusion of repeat purchase products in a competitive market: an agent-based simulation approach. Eur. J. Oper. Res. **245**(1), 157–167 (2015)

35. Ullrich, G.: The history of automated guided vehicle systems. In: Hollier, R.H. (ed.) Automated Guided Vehicle Systems, pp. 1–14. Springer, Berlin/Heidelberg (2015). https://doi.org/10.1007/978-3-662-44814-4_1
36. Vanhée, L., Dignum, F., Ferber, J.: Towards simulating the impact of national culture on organizations. In: MABS2013: 14th International Workshop on Multi-Agent-Based Simulation, Saint Paul, p. 12 (2013)
37. Vanhée, L., Borit, M., Santos, J.: Autonomous fishing vessels roving the seas: what multiagent systems have got to do with it. In: Proceedings of the 17th International Conference on Autonomous Agents and Multiagent Systems, pp. 1193–1197. International Foundation for Autonomous Agents and Multiagent Systems (2018)

Chapter 40
Conceptualising Artificial Anasazi with an Explicit Knowledge Representation and Population Model

Frederik Schaff

Abstract Using the case of the artificial Anasazi, it is shown how an explicit population model may be linked with an explicit network model. This network model is then used for the diffusion of knowledge, including the preservation of knowledge over generations. Knowledge is represented by an explicit model of cognition, in which each agent holds simulacra corresponding to entities in (virtual) reality, at the time when they were perceived. Heuristics exist to explore the artificial world in order to extend the cognitive model. In the selected example, knowledge items refer to land patches and attributes as well as fellow households, persons and settlements. The current status is work-in-progress. The original model without an explicit knowledge representation is already reimplemented and validated. The code is available via Github.

Keywords Agent-based modelling · Archaeology · Population model · Endogenous network · Explicit knowledge representation

40.1 Introduction

In current agent-based models (ABM), in archaeology individuals often do not interact directly. Instead, the environment acts as a mediator [5], similar to the *price* mechanism in neoclassical economics. Complete, although not necessarily perfect information is assumed regarding this environment – the economic correspondence would be prices and quantities. However, ABM offers the opportunity to deviate from this methodology and to emphasise agent myopia. In [10], the concept of

F. Schaff (✉)
Macroeconomics, Ruhr-Universität Bochum, Bochum, Germany

Project ReSoc, Leibniz Postgraduate School, Deutsches Begbau-Museum and Ruhr-Universität Bochum, Bochum, Germany
e-mail: frederik.schaff@rub.de

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_40

an explicit knowledge representation is developed for an economic setting. Here, this concept of *pure* agent-based computational economics (pACE) is applied to an archaeological setting.

40.2 Need for Advancing Artificial Anasazi

The artificial Anasazi project [1, 6, 7] serves as a good starting point: It has all the relevant aspects (geography, population model, decisions) and is assumed to be a closed system,[1] and it is assumed that a single activity, growing maize, is the only relevant economic aspect and that the choice of farm-site and settlement-site are the only variables at the agent‘s fingertip. The archaeological population record provides a benchmark and target for the model.[2] In addition, a very good data base exists regarding the geographical aspects of relevance for each year of the relevant time (800-1350 CE).[3] Based on the reimplementation by [9], which has a more accessible code than the original model (being written in NetLogo), we reimplement the model in LSD, a c++ based open-source cross-platform IDE.[4]

As has been noted before (e.g. in [9]), the original artificial Anasazi model is strongly driven by the environmental data, and some strong assumptions need to be in place in order to allow the model to produce its high fit with the archaeological data. Some of these assumptions relate to the population model. For example, there will only be new households if the environment has the capacity to support them. Other assumptions relate to the knowledge. For instance, it is implicitly assumed

[1]The Kayenta Anasazi lived in the Long House Valley in Arizona, USA. In previous modelling efforts, no external trade was assumed. To our knowledge the Anasazi had complex trade networks, and these might have been relevant to deal with local climate variance (cf., e.g. [11]). This might be a relevant future work. In fact, it is likely that the current model will not be sufficient to (completely) explain the archaeological record.

[2]This double-role is problematic for it can be argued that an equation with enough complexity can always be calibrated such that it replicates the desired time series. This can be countered by limiting the flexibility of the equation based on theoretic conditions. The explicit knowledge representation is a way to do so.

[3]The availability of drinking water, the hydrology that tells whether or not an area lies in a floodplain, a combined source called average Palmer Drought Severity Index that, together with some specific capacity of the soil that is distinguished in zones, provides input for the maize yield model.

[4]Laboratory for Simulation Development, see https://github.com/FrederikSchaff/Lsd/ for the current version used. You may find the Artificial Anasazi in the example model sections under GIS/Archaeology/ArtificialAnasazi. LSD is vividly developing recently and offers an easy to learn macro language that emphasises modular approaches. At the same time, the full power of modern c++ is used, both in terms of performance and versatility, allowing the complexity of models to be scaled with the experience of the user. Among the features are an automatic model documentation (hypertext and latex), several options for sensitivity analysis, parallelisation, cross-platform compatibility, an analysis of results module, export to R, integrated network capabilities and of course high performance, utilising, e.g. c++11.

that the agents know the exact productivity of each land patch in the Long House Valley (where the Anasazi lived), as well as whether or not there is (currently) a water source, etc. Although the relative proximity of the valley, 80 times 120 km, makes extensive exploration reasonable, it is nonetheless unreasonable that the exact attributes of each single patch of land are known at any given point to any agent. For example, agents know the current *potential* productivity of a patch of land even if they never farmed it. And even if that would be given, it is further assumed the Anasazi – each household individually and updated by the decision of the others – used a mechanism that allowed them to choose optimally among those 9,600 options (in model scale). Finally, a closer analysis of the code from [9] shows that there are some implementations in conflict with the written account in [1].[5] For example, households may farm at a spot that is not within 1600 m of a water source, and they may settle in a floodplain area. Also, households will not resettle if a water source vanishes or an area changes into a floodplain area.

Future plans: In difference to the original model studying the household-level, we choose the person-level. A number of persons (male and female, relatives) will form a household, and persons from a household can leave the household to form a new own household. Several households that reside in the same location form a settlement. We incorporate the simple statistical model used in [4] as a starting point to implement realistic age-distributions beyond the solution in [1].[6] Another mayor interest of the current study is the network formation process. It is assumed that some part of the network depends on the proximity to others (i.e. living in the same settlement), some part on the family (inheritance) and some part on a probabilistic interaction model. Furthermore, the heritage network is implemented in a way that it survives its nodes. The most relevant difference is, however, the representation of knowledge and the decision heuristics that work with this representation.

40.3 Explicit Knowledge Representation

For a better understanding of, e.g. settlement phenomena in archaeology, we suggest that each software agent embodies its own representation (i.e. model) of the geophysical reality but also social and other relevant reality. All its decisions should be based upon this internal model. This implies, for example, explicit mechanisms to search/explore unknown territory.

Because this internal representation is a construction based on the personal history of the agent, *sameness* becomes an affair of the internal agent‘s “mind”,

---

[5] Also noted by [9]. This is highlighted in an adjusted Version of Janssen’s model, included in the LSD reimplementation. In addition, the LSD reimplementation offers statistics on the number of cases where single assumptions are not kept.

[6] See https://github.com/FrederikSchaff/lsd-template-models/tree/master/Population for a template model already including [4] together with the potential to track family relations and a partner-search model based on [2, 3].

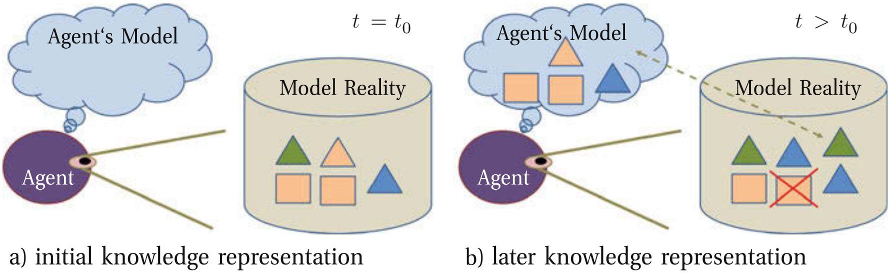


Fig. 40.1 Knowledge Representation in *p*ACE [10]. The *p*ACE agent may start with a blank model of the true state-space (the model world). By means of (heuristic) interaction with the environment, specific elements with their respective state at a specific time are perceived and remembered. The internal model may structurally differ from the external reality. Also, the state-space itself is allowed to change. (**a**) Initial knowledge representation. (**b**) Later knowledge representation

as demanded by [8, p. 371]. Figure 40.1 technically illustrates this idea from the concept of *pure* agent-based computational economics [10]. The items in the agent‘s model are purely individualised *simulacra* of the real items.

On the system level, we need to model agent communication and the general network of interaction to understand the propagation mechanisms for knowledge. The social network, as a consequence, should be endogenously dynamic. This can be handled by coupling the initialisation of a new person‘s social network position with the population model (i.e. ties to family) and basing the evolution of the network on an explicit interaction model. For example, chances to get new ties and hinder the degradation of existing ties correlate with the geographic proximity of the agents. Farming sites or water sources used by different settlements can then, for example, lead to a connection between those settlements. The transmission of knowledge is possible, whenever two connected persons visit the geographic area at the same time.

Some readers might be tempted to argue that a (standard) network approach instead of the simulacra approach is sufficient and more straightforward. This is, however, not the same as the approach just described. Here for example, each agent has its own internal composition of simulacra. The information held for each simulacra may or may not correspond to real attributes of the current status of the represented entity. In fact, it is in principle possible that the counterpart of a single simulacrum, what it represents, does not exist anymore, or never did (e.g., because the transmission process is faulty). This would then mean that, in order to have everything in one network, we needed to have multiple copies of the same real object; otherwise the handling of links (link attributes as "beliefs") would become very complicated. Another example: A single person that learns of another person X by talking to two different friends may not recognise that it is the same person referred to each time and will add two simulacra. On the other hand, a person that knows person Z in its role of a farmer, and now learns that she lives in the same settlement with person Z, will only add more information to the same simulacrum.

Also, cultural transmission, i.e. the passing of knowledge from older society members to younger members, including the ability of knowledge to survive its inventor, is explicitly considered by this methodology. A qualitative difference between first-hand knowledge (learned through own experience), second-hand knowledge (first-order information or knowledge passed to others)[7] and cultural knowledge (knowledge passed more than once, potentially between generations) can be made. For example, if migration within Long House Valley becomes necessary, the household may start looking for a suitable spot, where the grand-grand-parents once lived.

References

1. Axtell, R.L., Epstein, J.M., Dean, J.S., Gumerman, G.J., Swedlund, A.C., Harburger, J., Chakravarty, S., Hammond, R., Parker, J., Parker, M.: Population growth and collapse in a multiagent model of the Kayenta Anasazi in Long House Valley. Proc. Natl. Acad. Sci. **99**(Suppl 3), 7275–7279 (2002). https://doi.org/10.1073/pnas.092080799
2. Bijak, J., Hilton, J., Silverman, E., Cao, V.D.: Reforging the wedding ring: exploring a semi-artificial model of population for the United Kingdom with Gaussian process emulators. Demogr. Res. **29**, 729–766 (2013). https://doi.org/10.4054/DemRes.2013.29.27
3. Billari, F., Aparicio Diaz, B., Fent, T., Fürnkranz-Prskawetz, A.: The wedding-ring: an agent-based marriage model based on social interaction. Demogr. Res. **17**(3), 59–82 (2007). https://doi.org/10.4054/DemRes.2007.17.3
4. Boucekkine, R., de La Croix, D., Licandro, O.: Vintage human capital, demographic trends, and endogenous growth. J. Econ. Theory **104**(2), 340–375 (2002). https://doi.org/10.1006/jeth.2001.2854
5. Chliaoutakis, A., Chalkiadakis, G.: Agent-based modeling of ancient societies and their organization structure. Auton. Agent. Multi-Agent Syst. **30**(6), 1072–1116 (2016). https://doi.org/10.1007/s10458-016-9325-9
6. Dean, J., Gumerman, G., Epstein, J., Axtell, R., Swedlund, A., Parker, M., McCarroll, S.: Understanding Anasazi culture change through agent-based modeling: chapter 7. In: Kohler, T., Gumerman, G. (eds.) Dynamics in Human and Primate Societies, pp. 179–205. Oxford University Press, New York (2000)
7. Dean, J.S., Gumerman, G.J., Epstein, J.M., Axtell, R., Swedlund, A.C., Parker, M.T., McCarroll, S.: Understanding Anasazi Culture Change Through Agent-Based Modeling (1998). https://econpapers.repec.org/paper/wopsafiwp/98-10-094.htm
8. Georgescu-Roegen, N.: Analytical economics: issues and problems. Harvard University Press, Cambridge (1967)
9. Janssen, M.A.: Understanding artificial Anasazi. J. Artif. Soc. Soc. Simul. **12**(4), 13 (2009). http://jasss.soc.surrey.ac.uk/12/4/13.html
10. Schaff, F.: Pure agent-based computational economics of time, knowledge and procedural rationality with an application to environmental economics. Ph.D. thesis, FernUniversität in Hagen, Hagen (2016). http://nbn-resolving.org/urn:nbn:de:hbz:708-dh196
11. Wiener, J.B.: The mysterious ancient puebloan peoples (Anasazi): interview with: Stuart, D.E., Lund, J.M., DuBois, C. (2012). http://etc.ancient.eu/interviews/interview-the-ancient-anasazi-in-focus/

[7]In [10], fist-hand and second-hand information is considered for consumers sharing shopping experiences. At the same time, the first-hand experience of the same object can differ across individuals because it is evaluated subjectively based on the individuals history.

Chapter 41
Modeling Radicalization and Violent Extremism

F. LeRon Shults and Ross Gore

Abstract Given public anxiety about radicalization and violent extremism, it is not surprising that these topics have grabbed the attention of so many scholars in recent years. However, some have expressed concern over the fact that only a few studies in this relatively new field contain empirical data or systematic data analysis or develop causal models of the mechanisms generating these phenomena. We believe that computational modeling and simulation techniques can make a significant contribution to this scientific literature and eventually provide new tools for improving policy analysis. Here we briefly describe (1) an integrative theory of violent extremism proposed by Kruglanski and colleagues and (2) an agent-based model that instantiates this theory in a computational architecture.

Keywords Agent-based model · Radicalization · Violent extremism · Kruglanski

41.1 A "Needs, Narrative, and Networks" Theory

The work of Kruglanski and colleagues is based on social psychological research on the role that quest for significance (a microlevel variable) can play in motivating individuals toward radicalization and violent extremism [1]. However, their model also accounts for meso- and macro-level variables such as social processes and cultural ideologies that can promote and justify violent behaviors, which in turn provide some individuals with a significance gain. They propose that such behaviors are the result of the conjunction of three factors: a need that motivates one to (re)gain

F. L. Shults (✉)
Institute for Global Development and Planning, University of Agder, Kristiansand, Norway

Center for Modeling Social Systems at NORCE, Kristiansand, Norway
e-mail: leron.shults@uia.no

R. Gore
Virginia Modeling, Analysis and Simulation Center, Old Dominion University, Suffolk, VA, USA

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_41

personal significance, the availability of a narrative or ideology that can justify the behavior, and a social network whose group dynamics lead an individual to embrace the ideology [2]. Finally, their theory also recognizes that an individual must have the ability (subjective and objective capacity) to carry out the extreme behavior.

Extremism is "motivated deviance from general behavioral social norms," which occurs as a result of "a shift from a balanced satisfaction of basic human needs afforded by moderation to a motivational imbalance wherein a given need dominates the others" [3, p., 217]. The authors apply this general psychological theory of extremism to the special case of violent extremism. All humans need significance, although this variable is differentially distributed among individuals in a population. Engaging in extreme behaviors takes a great deal of energy, and most people maintain their motivational balance by engaging in moderate behaviors that fulfill a variety of needs. However, motivational imbalance can be triggered by traumatic experiences or cultural threats, which can increase the quest for significance. If that need is sufficiently high, and an individual becomes entangled in radicalizing social networks whose group dynamics lead him or her to accept a violence-justifying ideological narrative, and that individual has the ability to perform such acts, he or she is more likely to engage in extremist behaviors.

Elsewhere, our research team has developed models directly or indirectly related to these themes, including simulations of the dynamics involved in terror management systems [4], mutually escalating religious violence [5], and existential security and secularization [6]. Some of these models included variables related to those discussed by Kruglanski and his colleagues, including identity fusion, sacred values, and concerns about threat or existential anxiety. The model outlined below implements the "needs, narrative, and network" theory within a computational architecture that incorporates the relevant micro-, meso-, and macro-level variables. This is our first attempt to demonstrate the value of computational modeling and simulation techniques for predicting and preventing radicalization [7, 8].

41.2 A Computational Model of Radicalization and Violent Extremism (RAVE)

Each agent in RAVE has a quest for significance level (QS), a level of motivational imbalance (MI), a cultural threat threshold, an ideological narrative violence level (INV), a threshold specifying a willingness to commit violence (VE), and a minimum, mean and maximum ability to commit violence. In addition, each agent is assigned a social network, which reflects its set of active social relationships. The number and weight of the links within the network is constructed by an algorithm that mirrors real human social networks, in which individuals tend to have ∼135 active relationships [9]. Based on homophily indicators derived from research in social anthropology, we tie strong relationships within the network to higher levels of emotional closeness. In our implementation, emotional closeness is also correlated with an agent's INV level. The degree to which the INV of an agent is correlated to the emotional closeness of its network relations is determined by the parameter INV-homophily (INV-H).

The model environment produces cultural threats every time step. Every threat has an intensity determined by a triangular distribution with a minimum, mode, and maximum parameter. The values of these parameters can range from 0 (no intensity) to 1 (maximum intensity). Each agent also has a threshold ranging from 0 (no threshold) to 1 (maximum threshold) that determines how intense a cultural threat needs to be for them to perceive it. The perception of cultural threats causes agents to increase their motivational imbalance. By contrast, not perceiving threats decreases an agent's level of motivational imbalance. Within each agent, a motivational imbalance level is represented as a real number.

Figure 41.1 summarizes the interactions and decisions made by entities within the model at each time step. The left side of Fig. 41.1 represents those actions that occur if an agent's MI level is below the agent's QS threshold. In this case, cultural threats are created for the agent. The intensity of each threat is determined by sampling the triangular distribution defined by the model-level parameters. Once

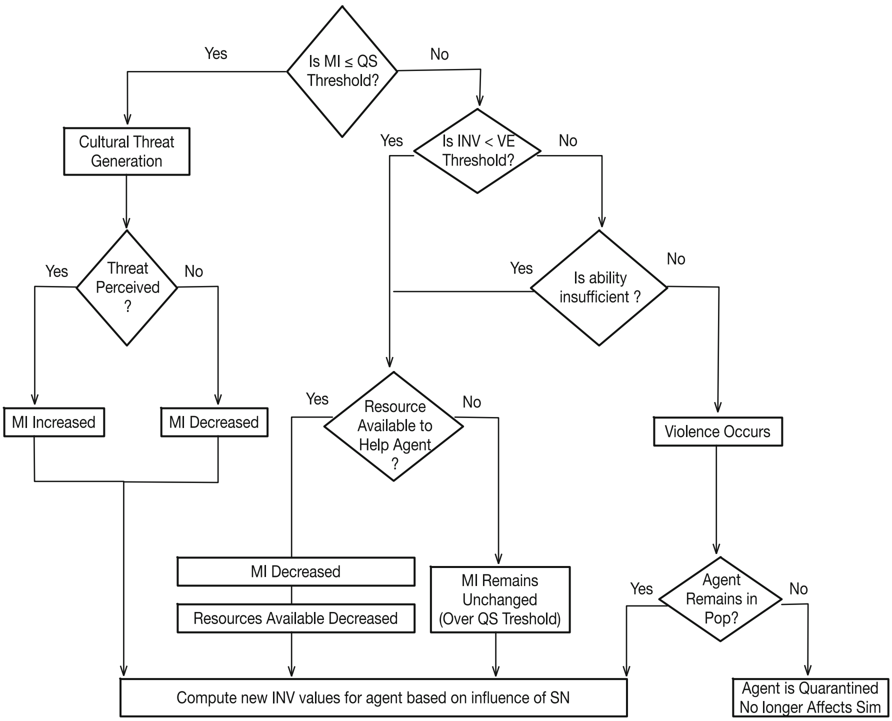


Fig. 41.1 Decisions made by entities within RAVE at each time step

an agent tests for hazard perception, it interacts with its social network, and its INV level is updated using a Rescorla–Wagner formula for classical conditioning [10]. In our implementation, this formula is $\mathrm{MI}_{\mathrm{current}} + [\alpha\beta(\lambda - \mathrm{MI}_{\mathrm{current}})]$. In the formula, when hazards are perceived, the rate of the stimuli (β) is set to 1, and the association value (λ) is set to 1. These parameters reflect the increase of the MI level of an agent in response to incoming cultural threats. Conversely, when hazards are not perceived, the rate of the stimuli (β) is set to 0.1 and the association value (λ) is set to 0. This reflects the decay of the agent's MI without the presence of the stimulus. The value of α remains constant in both cases.

The right side of Fig. 41.1 depicts actions that are initiated if an agent's MI exceeds its quest for significance threshold (QS). When this occurs, the agent checks its INV level (the current level of violence associated with the agent's ideological narrative), which is influenced by the agent's social network. If an agent's INV exceeds its willingness to commit violence (VE), then the ability of the agent to commit violence is determined by sampling the triangular distribution with the minimum, mode, and maximum parameter. The uncertainty within these parameters reflects the variance in an agent's capacity to commit violence over time. If the agent's ability is over the specified threshold for the model, then the agent commits a violent act. When an agent commits a violent act, then the simulation checks to see if the agent will be removed from the population. This check is performed by sampling the triangular distribution (described above) and testing whether the sampled number is above the removal after violent act threshold; if so, the agent is removed from the simulation and no longer affects the run. If not, the agent remains in the simulation with a motivational imbalance that is above the QS threshold, and its INV level is updated based on its interaction with its social network.

If (1) an agent's INV does not exceed their willingness to commit violence or (2) an agent's ability to commit violence is less than or equal to the specified threshold, then the agent does not commit a violent act. In either case, the agent checks to see whether there are resources in the environment that could provide some other way to fulfill the quest for significance (e.g., community programs), thereby decreasing motivational imbalance level. If resources are available in the environment, then (1) the motivational imbalance of the agent is decreased, and (2) the number of resources available in the environment for all future agents is decreased by one. If resources are not available, the motivational imbalance of the agent remains over the QS threshold. In either case, the agent then interacts with its social network and updates its INV level. In this interaction, the INV variable within the agent is influenced by the INV value of other agents within their social network. It is important to note that Fig. 41.1 shows that the influence of the social network is exerted every time step of the model on an agent even if she/he does not perceive a cultural threat. The only way an agent can avoid being influenced by their social network during a time step is if the agent is quarantined.

The extent to which the variable is influenced is determined by a time-dependent weighted average. Given a matrix IN that includes an entry for the influence of each of the N agents on every other agent, the total influence exerted on it is computed using the equation described next. Total Influence and a set A that includes all agents enable us to define $\text{TotalInfluence}_i = \sum_{\substack{j=1 \\ i \neq j}}^{N} \text{IN}_{i,j}$. Total Influence and a set A that includes all agents enable us to define $A_{\text{SN}-\text{INV}_{t,i}}$. Set A contains the value of INV, at each time step t, for each agent j, throughout the simulation ($A_{\text{INV}_{t,j}}$). $A_{\text{SN}-\text{INV}_{t,i}}$ is the influence exerted on agent i by his/her social network (SN) for a given variable v at time t. Formally, it is $A_{\text{SN}-\text{INV}_{t,i}} = \sum_{\substack{j=1 \\ i \neq j}}^{N} \frac{A_{\text{INV}_{t,j}} * \text{IN}_{i,j}}{\text{TotalInfluence}_i}$. An agent combines the value of the INV variable from their social network with the agent's existing value for the respective variable using the Cobb–Douglas function [11]. We employ this function because it is an established, flexible, and widely used method to aggregate the influence of the environment with the existing value of a variable though the parameter β. Formally, this combination is computed as $A_{\text{INV}_{t+1,i}} = {A_{\text{INV}_{t,i}}}^{\beta} * {A_{\text{SN}-\text{INV}_{t,i}}}^{1-\beta}$.

The model has several important assumptions. First, note that resources in the model cannot be restored or increased. This reflects the assumption that the model will be used for planning or policy decisions for a fixed period of time. Thus, the model explores how few resources need to be budgeted for a fixed period of time (e.g., a year) to minimize the number of violent actions in the artificial society. Second, a violent action that results in quarantine is assumed to have direct consequences only on the extremist agent who performed the action. This excludes effects on the extremist's network and on the rest of the population. As a result, this model cannot capture some other dynamics that may play a role in radicalization processes, such as imitation. Third, the model assumes that the cultural threats an agent perceives are independent of his or her social network. This assumption precludes cultural threats that stem from one's social network.

In future work, we plan to construct simulation experiments within this artificial society that will allow us to explore the potential impact of various policies for mitigating radicalization and extremism [12, 13].

Acknowledgments The authors are members of the "Modeling Religion in Norway" (MODRN) project, which is funded by The Research Council of Norway (grant #250449).

References

1. Kruglanski, A.W., Gelfand, M.J., Bélanger, J.J., Sheveland, A., Hetiarachchi, M., Gunaratna, R.: The psychology of radicalization and deradicalization: how significance quest impacts violent extremism. Polit. Psychol. **35**, 69–93 (2014)
2. Webber, D., et al.: The road to extremism: field and experimental evidence that significance loss-induced need for closure fosters radicalization. J. Pers. Soc. Psychol. **114**(2), 270–285 (2017)
3. Kruglanski, A.W., Jasko, K., Chernikova, M., Dugas, M., Webber, D.: To the fringe and back: violent extremism and the psychology of deviance. Am. Psychol. **72**(3), 217–230 (2017)

4. Shults, F.L., Lane, J.E., Diallo, S., Lynch, C., Wildman, W.J., Gore, R.: Modeling terror management theory: computer simulations of the impact of mortality salience on religiosity. Relig. Brain Behav. **8**(1), 77–100 (2018)
5. Shults, F.L., Gore, R., Wildman, W.J., Lynch, C.J., Lane, J.E., Toft, M.D.: A generative model of the mutual escalation of anxiety between religious groups. J. Artif. Soc. Soc. Simul. **21**(4), 1–24 (2018)
6. Gore, R., Lemos, C., Shults, F.L., Wildman, W.J.: Forecasting changes in religiosity and existential security with an agent-based model. J. Artif. Soc. Soc. Simul. **21**, 1–31 (2018)
7. Shults, F.L. Can we predict and prevent religious radicalization?. In: Øverland G. (ed.) Processes of Violent Extremism in the 21st Century: International and Interdisciplinary Perspectives. Cambridge Scholars Press, Cambridge (2018), pp. 45-71.
8. Shults, F.L.: Practicing Safe Sects: Religious Reproduction in Scientific and Philosophical Perspective. Brill Academic, Leiden (2018)
9. Conti, M., Passarella, A., Pezzoni, F. A model for the generation of social network graphs. In: World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011 IEEE, 2011, pp. 1–6
10. Rescorla, R.A., Wagner, A.R., et al.: A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. Class. Cond. II Curr. Res. Theory. **2**, 64–99 (1972)
11. Cobb, C., Douglas, P.: A theory of production. Am. Econ. Rev. **18**(1), 139–165
12. Christmann, K. Preventing Religious Radicalisation and Violent Extremism: A Systematic Review of the Research Literature. Applied Criminologyv Centre, University of Huddersfield, Huddersfield, UK (2012)
13. Neumann, P., Kleinmann, S.: How rigorous is radicalization research? Democr. Secur. **9**(4), 360–382 (Sep. 2013)

Chapter 42
The Artificial Society Analytics Platform

F. LeRon Shults, Wesley J. Wildman, Saikou Diallo, Ivan Puga-Gonzalez, and David Voas

Abstract Social simulation routinely involves the construction of artificial societies and agents within such societies. Currently there is insufficient discussion of best practices regarding the construction process. This chapter introduces the artificial society analytics platform (ASAP) as a way to spark discussion of best practices. ASAP is designed to be an extensible architecture capable of functioning as the core of many different types of inquiries into social dynamics. Here we describe ASAP, focusing on design decisions in several key areas, thereby exposing our assumptions and reasoning to critical scrutiny, hoping for discussion that can advance debate over best practices in artificial society construction. The five design decisions are related to agent characteristics, neighborhood interactions, evaluating agent credibility, agent marriage, and heritability of personality.

Keywords Artificial society · Agent-based model · Design decisions

F. L. Shults (✉)
Institute for Global Development and Planning, University of Agder, Kristiansand, Norway

Center for Modeling Social Systems at NORCE, Kristiansand, Norway
e-mail: leron.shults@uia.no

W. J. Wildman
Center for Mind and Culture, Boston, MA, USA

Boston University, Boston, MA, USA
e-mail: wwildman@bu.edu

S. Diallo
Virginia Modeling, Analysis and Simulation Center, Norfolk, VA, USA

I. Puga-Gonzalez
Institute for Global Development and Planning, University of Agder, Kristiansand, Norway
e-mail: ivan.puga-gonzalez@uia.no

D. Voas
University College London, London, UK
e-mail: d.voas@ucl.ac.uk

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_42

42.1 Introduction

This chapter introduces the artificial society analytics platform (ASAP), a computational model designed to function as an extensible artificial society for studying social life in developed western cities. Our purpose here is to discuss some of the design challenges we faced when constructing ASAP. By surfacing these challenges and explaining our design decisions, we hope to foster dialogue about best practices in building artificial societies. In other words, we are addressing the theme of the 2018 SSC by "looking in the mirror" – taking a hard look at how well we are doing in constructing realistic and functional artificial social worlds. Societies are so complex that their intricacies may seem to defy computational modeling. However, we believe that the construction of sophisticated artificial societies is becoming increasingly feasible – that is, both computationally tractable and materially useful for scientific research.

A wide variety of general agent-based models have been developed over the years as scholars in this field have made advances in simulating the emergence of macro-level societal phenomena from microlevel interactions [1–9]. We believe the field of social simulation can mature even more quickly if we engage in sustained, self-critical discussions about how well we are modeling critical aspects of social complexity.

We designed ASAP to handle a variety of specialized inquiries related to the evaluation of scientific hypotheses about – and policies for promoting – healthy social equilibria within urban areas in the developed western world. The platform features "worldview clubs," which may be either religious or secular; this is intended to facilitate the exploration of hypotheses about the dynamics of group life within and between religious and nonreligious groups. It includes a majority (host) population group and a minority (immigrant) population group and individual level variables such as outgroup suspicion, ingroup support, and shared norms, which make it suitable for extension to models about immigrant integration. Agents in ASAP are distributed in different neighborhoods and linked to job locations. They meet in a variety of networks, influence one another's worldviews, get educated, seek employment, look for compatible marriage partners, and reproduce, age, and die. These features render ASAP useful for certain types of policy modeling, estimating cost-effectiveness of policy proposals, and informing debates among policy professionals.

In the latter part of this chapter, we describe the agents, agent interactions, and parameters within the computational architecture of ASAP. We conclude with reflections on our experience of looking in the mirror. However, we begin by offering our rationale for some of the design decisions we had to make in constructing an artificial society that could be extended to multiple case studies. Our goal is to foster debate about such decisions among those invested in promoting and improving social simulation methodologies.

42.2 Design Decisions for Discussion

42.2.1 Agent Characteristics

The first series of design decisions for scrutiny deal with the definitions of agent variables (which are described in more detail in Sect. 42.3). Without complex agents in play, an artificial society will be incapable of expressing social phenomena of interest. If agents are too complex, we risk losing cognitive control over the model and creating a computational monster too slow to be useful. Finding the sweet spot depends on the specific inquiry for which the virtual society is being developed and on available computational resources. We plan to use ASAP to answer a variety of questions, and so it needs to be extensible in the relevant ways. Our goal was to identify an ideal set of core agent characteristics that could be used in almost all specific applications.

We settled on the agent variables portrayed in Fig. 42.1 for several reasons. The rationale for the *demographic* variables is that they seem to be those minimally required to simulate almost any interesting social dynamics. These demographic

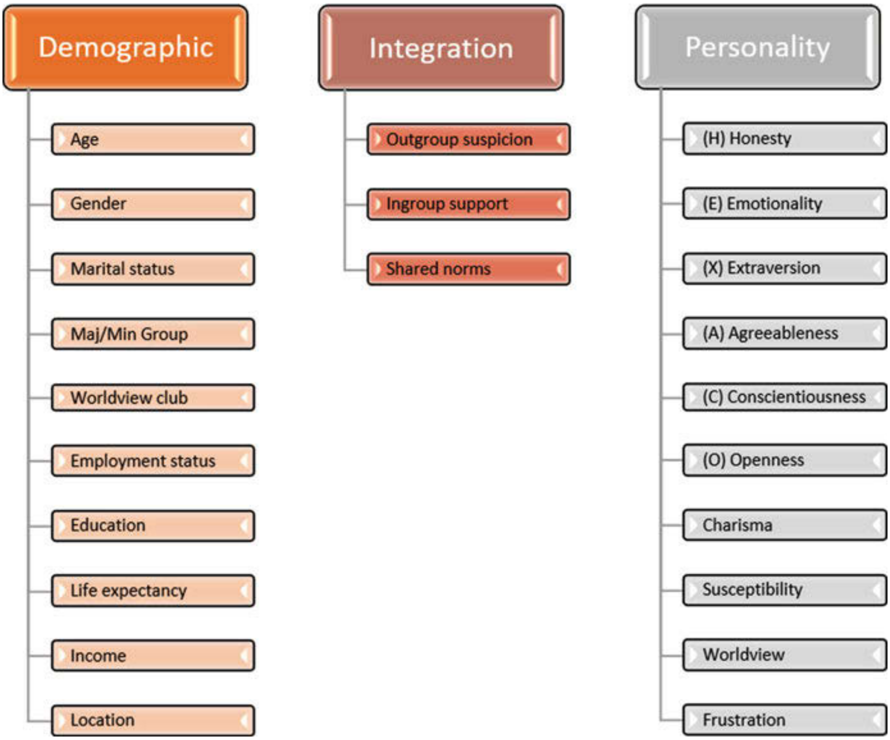


Fig. 42.1 Three types of agent variables in ASAP

variables frequently interact in ways that are useful for policy modeling. For example, policies aimed at increasing immigrant employment can have negative effects on majority employment rate in western, urban societies of the sort ASAP is designed to simulate.

Although other sociological theories were taken into account in framing the *integration* variables, the computational architecture was strongly informed by the work of sociologist of religion David Voas [10, 11]. These variables combine with some of the demographic variables to help capture the difference between three distinct dimensions or types of integration: (1) *structural* integration, characterized by equality of opportunity in education, employment, housing, civil rights, and civic participation; (2) *social* integration, defined as interaction between members of different groups in ways that range from the superficial (brief impersonal encounters, e.g., in commercial transactions) to the deeply personal (close friendships and intimate relationships); and (3) *cultural* integration, which involves shared norms, values, worldviews, and cultural capital.

We adopted the HEXACO framework [12] for *personality* variables, distributing them in the artificial population in ways that reflect the real world. We preferred HEXACO because it extends the Big Five framework [13] by adding honesty/humility, which is an important factor in some interpersonal interactions. Incorporating personality enables us to enrich the representation of several social dynamics, including shifts in agent worldviews and switching in religious/secular worldview club affiliation, since these worldview-related variables are associated with personality (e.g., religious individuals, especially those in religious worldview clubs such as churches or synagogues or temples, tend to be higher in conscientiousness and lower in openness than the population as a whole).

Beyond HEXACO variables, we settled on three personality features critical for interpersonal interactions. The intensity of the effect of an interaction between ego and alter agents is amplified or muted by the *susceptibility* of the ego and the *charisma* of the alter. The *worldview* variable enables us to characterize an agent's religious (supernaturalist) or secular (naturalist) way of thinking and the way it changes during personal interactions, which is crucial for studying worldview club affiliation, disaffiliation, and reaffiliation dynamics.

42.2.2 Neighborhood Interactions

Each agent may engage in several sorts of interactions each week. The frequency of salient interactions capable of impacting ego agent variables is an interesting question. We settled on weekly interactions partly because interactions that occur more frequently are rarely significant and partly to simplify computational load. We know of no quantitative data to answer the frequency question; hence, we relied on the intuition of our subject-matter experts (SMEs). To focus debate, here we discuss interactions in an agent's neighborhood (around the agent's place of residence), which are particularly important when ASAP is used to study immigrant integration.

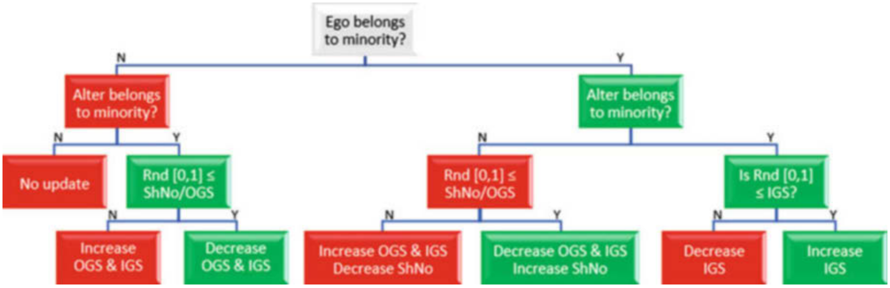


Fig. 42.2 Dyadic neighborhood interactions in ASAP. Ego = agent possibly changing, alter = ego's interaction partner, Rnd [0,1] = random number between 0 and 1, IGS = ingroup support, OGS = outgroup suspicion, ShNo = shared norms

Figure 42.2 portrays the decision tree for an ego agent interacting with an alter agent in ego's neighborhood (this is an enlarged version of the matching part of Fig. 42.5 below). We made several assumptions. First, we assumed that minority/majority interactions must be analyzed separately from same-group interactions. Second, we assumed that effects of interactions would depend on a stochastic process in which the relative importance of shared norms and outgroup suspicion govern how well an interaction is likely to go. Third, we assumed that the impact of neighborhood interactions is adequately captured by changes in the three integration agent variables. Fourth, we assumed no other considerations were important enough to include. Each of these assumptions was based on SME intuition, and each is worth discussing among those who build virtual societies.

42.2.3 Evaluating Agent Credibility During Personal Interactions

We also use ASAP to study secularization, a process facilitated by a lack of religious credibility enhancing displays (CREDs) within a population [14, 15]. Research in this area shows that individuals are more likely to continue believing in the gods of the religious clubs in which they were raised, and affiliating with those clubs, if they encountered during childhood costly displays of belief within their families of origin and religious contexts. When there are inadequate displays to enhance credibility, a population is more likely to become secular, especially when governmental institutions satisfy basic human needs without dependence on religious institutions. This is an example of empirical findings and theoretical developments in the relevant sciences informing our design decisions. We assume that the cohesion of groups whose identity is connected to secular ideology is also partially dependent on the presence of sufficient CREDs in such groups.

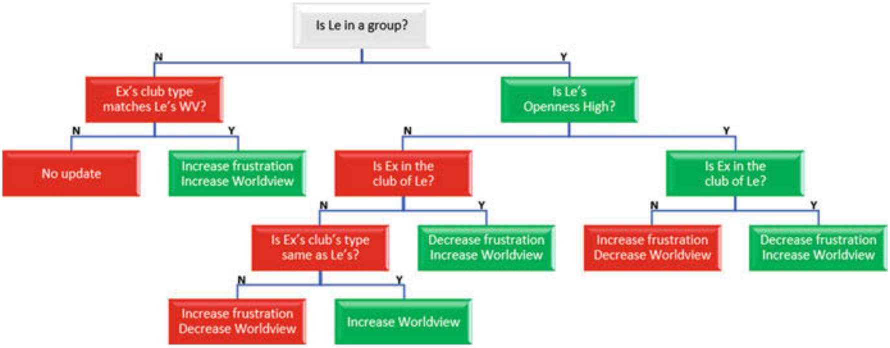


Fig. 42.3 Decision tree for evaluating credibility of behaviors in ASAP. Le = learner (ego) agent, Ex = exemplar (alter) agent, club type is Boolean (0 = secular, 1 = religious)

We conceived interactions relevant to CRED evaluation in terms of a conceptual framework from epidemiology of representations: the ego agent is a *learner* whose variables are subject to change and the alter agent is an *exemplar* who potentially impacts the learner. The worldview interaction is distinct from the six depicted in Fig. 42.2. We present the learner agent's decision tree for evaluating an exemplar's potential CRED in Fig. 42.3. The learner variables that may change are worldview, which varies continuously between 0 (fully secular and naturalist) and 1 (fully religious and supernaturalist), and frustration, which varies continuously between 0 (not at all frustrated with the club ego is in or with being in no club) and 1 (extremely frustrated with ego's club or with being in no club).

When frustration passes a threshold, ego will act in one of three ways: (1) joining the club of the last club-affiliated person who impacted ego positively; (2) switching from one club to another, joining the club of the last club-affiliated person who impacted ego positively; or (3) leaving ego's current club and having no worldview club. After acting, ego's frustration variable drops dramatically. The worldview club joined tends to be compatible with ego's worldview. It is possible for a secular-leaning ego agent to affiliate with a religious group or vice versa, but a person with an extreme worldview value would never join a mismatched worldview club.

Critical to the dynamics of worldview change and worldview club affiliation is the ego-learner agent's openness. When ego is in a club, the openness personality characteristic plays an important role in assessing the impact of a CRED from someone in a different club. If ego is low in openness, a CRED from a member of a competitor club will threaten ego, increasing frustration and changing worldview – we call this the pluralism effect, and it acts on low-openness agents. By contrast, if ego is high in openness, a CRED from a member of a competitor club will not trigger the pluralism effect but rather increase worldview confidence and decrease ego's frustration with ego's current group. These dynamics depend on findings in social-psychological and personality research on religious pluralism.

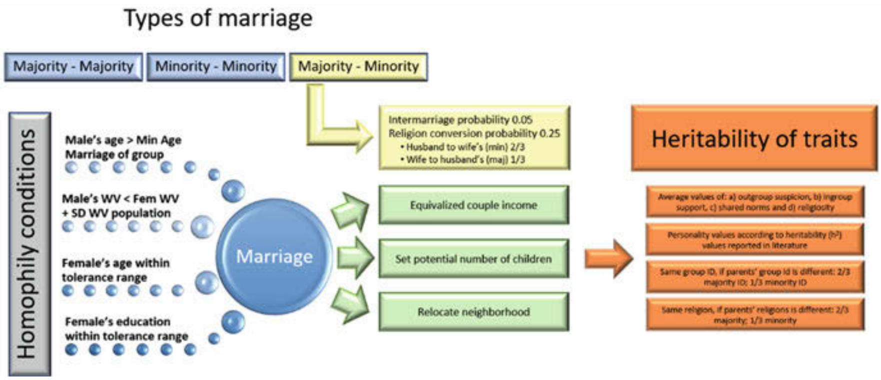


Fig. 42.4 The marriage process. WV = worldview values, SD = standard deviation. Tolerance ranges are set in model parameters

42.2.4 Agent Marriage and Homophily Constraints

Marriage plays an important role in shaping all human interactions and especially in the integration of cultures. We therefore tried to make design decisions that would capture the most relevant causes and effects related to marriage that bear on the social dynamics of urban western societies.

Many traits related to religion and personality affect social interactions and are also heritable; hence marriage and producing biological offspring are important conditions for the spread of these traits through a population. Moreover, our SMEs guided us through what social theorists know about homophily in the process of selecting a marriage partner. Figure 42.4 presents the homophily constraints on the left, the variable changes upon marriage in the center, and the handling of heritable traits in offspring on the right (discussed below). Regarding homophily conditions, we selected the four considerations our SMEs thought were most important on average; these assumptions should be debated.

42.2.5 Heritability of Personality

ASAP posits that personality (instrumentalized primarily using the HEXACO six-factor system) informs most relevant personal characteristics and, being significantly heritable, is the principal vehicle for the transmission of traits across generations. We derive heritability factors from research in twin studies [16, 17]. We round heritability factors to the nearest one-sixth in Table 42.1. Heritability of personality is critical for social dynamics. For instance, we know from twin studies that religiosity, conservatism, and authoritarianism are significantly co-heritable

Table 42.1 Children inherit personality traits from parents with heritability (h^2) values, which indicates the extent of genetic (as against environmental) influence

Personality trait	Heritability (h^2)
H – Honesty	0.33
E – Emotionality	0.33
X – Extraversion	0.50
A – Agreeableness	0.33
C – Conscientiousness	0.50
O – Openness	0.67

and that this can be captured through the transmission of personality traits across generations because religiosity, conservatism, and authoritarianism have strong personality correlations particularly with openness and conscientiousness.

42.3 Agent Variables in ASAP

The simulated agents in ASAP inhabit a virtual world where they attend school, get hired and fired, get married, and reproduce (for a full list of variables, see Table 42.2). Agents are categorized into majority or minority *groups* depending on their family of origin. They have variables related to their *demographic* features (age, group, education, employment, etc.), their level of *integration* into society (outgroup suspicion, ingroup support, and shared norms), their personal *worldview* (ranging from religious supernaturalism to secular naturalism), and their worldview *club* identification (religious club, secular club, no club). The values and settings for the demographic variables are based on available data and/or subject matter expert assessment. For example, we believe that immigrants in the contexts we want to simulate will tend to be more highly educated due to factors such as immigrant selectivity and the high value immigrant parents place on education as a means of social mobility. This is the sort of assumption that ought to be debated in the construction of artificial societies.

Worldview clubs are membership organizations that exist to support people having specific types of worldview and to advance those worldviews; when an agent belongs to a worldview club, the agent's personal worldview variable tends to match the worldview of the club. Agents have personality (using the six HEXACO personality factors and several other personality features), memories of salient interpersonal encounters, the ability to learn from others, and the power to evaluate the sincerity and consistency or hypocrisy of those around them. In the process of interpersonal exchange, agents may change their worldviews and their club-affiliation identities.

On initialization of a simulation run, agents are assigned variables drawn from suitable distributions that may vary according to whether they are in the majority or minority. They attend school until they are at least 16 years old; thereafter the number of years of further education they receive varies. After finishing school,

Table 42.2 Variables held by agents in the ASAP platform

Variable	Description	Values	[Min,Max]
Status	Status of agent	Student/employed/unemployed	NA
Life expectancy Maj	Potential lifespan for majority	Triangular (min,max,Mode)	[65,95,80]
Life expectancy min	Potential lifespan for minority	Triangular (min,max,Mode)	[45,85,65]
Total education Maj	Years of education	Min + Normal (μ;σ^2)	[10, 20]
Total education min	Years of education	Min + Normal (μ;σ^2)	[12, 20]
Suspicion Maj	Level of suspicion toward min	Normal (μ,σ^2) or average of parents	[0,1]
Suspicion min	Level of suspicion toward Maj	Normal (μ,σ^2) or average of parents	[0,1]
Group support Maj	Level of Maj ingroup support	Normal (μ,σ^2) or average of parents	[0,1]
Group support min	Level of min ingroup support	Normal (μ,σ^2) or average of parents	[0,1]
Shared norms	Degree of sharing cultural norms	Normal (μ,σ^2) or average of parents	[0,1]
Number of children	Likelihood of having children	$0 = 16\%$, $1 = 18\%$, $2 = 41\%$, $3 = 18\%$, $4 = 7\%$	[0,4]
Level of authority	Authority given by employment	Uniform [0,1]	[0,1]
Worldview	Worldview identification	Normal (μ,σ^2) or average of parents	[0,1]
(H)honesty	Personality trait	Normal (μ,σ^2) or inherit from parents	[0,1]
(E)emotionality	Personality trait	Normal (μ,σ^2) or inherit from parents	[0,1]
(X)extraversion	Personality trait	Normal (μ,σ^2) or inherit from parents	[0,1]
(A)agreeableness	Personality trait	Normal (μ,σ^2) or inherit from parents	[0,1]
(C)conscientiousness	Personality trait	Normal (μ,σ^2) or inherit from parents	[0,1]
(O)openness	Personality trait	Normal (μ,σ^2) or inherit from parents	[0,1]
Charisma	Personality trait	Normal (μ,σ^2)	[0,1]
Susceptibility	Personality trait	Normal (μ,σ^2)	[0,1]

Triangular = values from triangular distribution (with Min = minimum, Max = maximum, Mode = mode). Normal = values from a normal distribution (with mean μ, standard deviation σ^2). Uniform = values from a uniform distribution (over an interval) *Maj* majority, *Min* minority, *NA* not applicable

agents attempt to move into the work force; the likelihood of agents getting employed depends on their sex and majority or minority group classification. Agents die with a certain probability or if they reach a natural limit derived from a longevity distribution. They have a chance of getting married once they reach the relevant age threshold. Agents tend to marry agents of their own group, but mixed marriages (between majority and minority groups) are also possible. To get married, agents must satisfy age, education, and worldview compatibility conditions related to their potential partner (Fig. 42.4 above). Once married, agents may have children; newly born agents inherit the personality traits of one of their parents (randomly chosen) and the average value of the integration variables (outgroup suspicion, ingroup support, and shared norms). Other demographic variables are assigned according to parameter values specified in the model.

42.4 Agent Interactions

Agents have up to six different kinds of social interactions on a weekly basis (see Fig. 42.5). To interact, agents must be at least 12 years old. A work interaction requires the agent to be employed. Agents interact with others within their family, work, neighborhood, online, offline, and impersonal social networks. Family networks consist of the father and mother, neighborhood networks are all agents in the same neighborhood as the ego agent, online networks are two agents selected at random from the entire population, work networks are all agents working at the same job location, and impersonal networks are agents within *interaction radius* distance from ego. Offline social networks are stochastic with the probability of being someone else's alter agent inversely proportional to the spatial distance between ego and alter. Every week, an alter agent from each network is selected at random, and an interaction with ego occurs. These interactions result in positive, negative, or neutral outcomes, which increase, decrease, or leave equal ego agent variables related to integration: outgroup suspicion, ingroup support, and shared norms.

Majority/majority interactions produce no changes. A minority agent evaluates interactions with other minority agents based on the average degree of ingroup support in the minority group: if average minority ingroup support is higher than a random number between [0,1], then the result of the interaction is positive (negative otherwise). The outcome of minority/majority interactions depends on the ratio between average agreement level (shared norms) and average outgroup suspicion in the entire population: if the average agreement level is higher than that of the outgroup suspicion, the interaction will likely be positive (negative otherwise). There is a stochastic element, which is why random numbers appear in Fig. 42.5.

In work and impersonal interactions, average outgroup suspicion is multiplied by *AntiDis* (AntiDiscrimination) and *MinFr* (Minority Friendly), respectively. *AntiDis* and *MinFr* represent the strength of antidiscrimination laws and multicultural behavior present at ego's job or in ego's impersonal network. Multiplying outgroup suspicion by these variables increases the likelihood of a positive interaction. In impersonal minority/majority interactions, the level of authority of the alter impacts the update of the ego's integration variables. The rationale is that a positive (or neg-

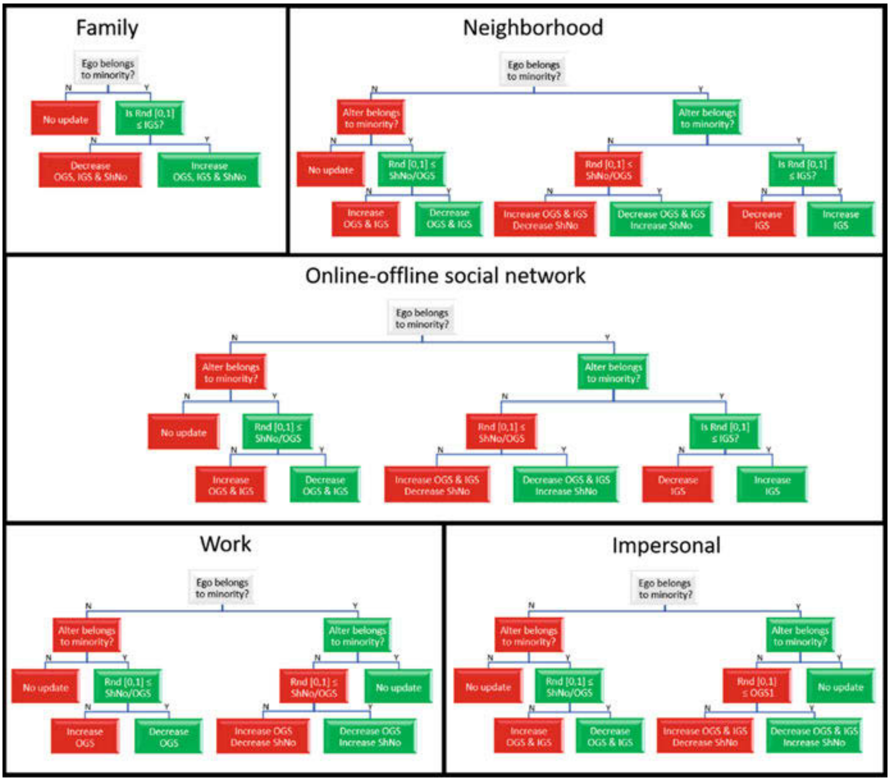


Fig. 42.5 Weekly interactions in the model and their effect on the agents' integration variables. OGS = outgroup suspicion, IGS = ingroup support, ShNo = average shared norms for agent's group

ative) interaction in an unimportant encounter (measured by the level of authority of the alter) will have a lower impact than one with a more powerful alter. In family, neighborhood, and offline interactions, women in minority/minority interactions are more vulnerable and thus more likely (10%) than men to rate interactions positively (because of pressure from family, friends, and/or neighborhood); in the model, this is represented by subtracting 0.1 from the random number [0,1].

The update of values after an interaction is shown in Table 42.3. The ego agent increases (or decreases) its current integration variable value by adding (or subtracting) a random value drawn from a normal distribution with mean and standard deviation obtained from the values of the alter agent's group. In the case of family interactions, the random value is drawn from a normal distribution with mean and standard deviation obtained from ego's parents' values. Note that these variables range from 0 to 1; if the updated value goes beyond this range, then the variable is set to its maximum (1) or minimum (0) value.

Table 42.3 Update of integration variables (outgroup suspicion, ingroup support, shared norms) after positive interactions

Interaction	Updated value	μ,σ^2 from
Family	CV + Normal (μ,σ^2)	Parents
Neighborhood	CV + Normal (μ,σ^2)	Agent's group
Online	CV + Normal (μ,σ^2)	Agent's group
Offline	CV + Normal (μ,σ^2)	Agent's group
Work	CV + Normal (μ,σ^2)	Agent's group
Impersonal	CV + (authority impact*Alter's authority level) + Normal (μ,σ^2)	Agent's group

The update of negative interactions is the same, but the plus (+) sign after current value is changed to minus (−). CV = current value, Normal = values drawn from a normal distribution (with mean μ, standard deviation σ^2), Authority Impact is a parameter

After a year of social interactions (52 weeks), agents (a) die if they reach their life expectancy; (b) may die with a certain probability (*Death probability*, see Table 42.4); (c) may get married and have children; (d) update their online/offline social networks; (e) if in the work force, may obtain/lose jobs; (f) if students, increase their education level; and (g) if completing their education, enter the work force and may obtain jobs.

42.5 ASAP Parameters

Parameters in ASAP are fixed for a given simulation run and can be varied to explore model dynamics. Table 42.4 lists the parameters. Since we are not reporting on experiments, we furnish this as contextual information for explaining our design decisions.

42.6 Verification, Validation, Calibration, Computability

ASAP has been built and we have verified that it conforms to requirements. We are in the process of working with subject-matter experts to validate it against data. One validation method is examining a matrix of parameter (input) and variable (output) values drawn from a parameter sweep, looking for plausible and implausible correlations. We have not yet attempted to calibrate ASAP to specific contemporary western cities.

ASAP can be run on a standard 64-bit windows platform with 64 GB of RAM and an Intel® Xeon® processor at 2.60 GHz. The model is built on top of the AnyLogic® platform which is a java-based environment for developing hybrid models (discrete-event, system dynamics, and agent based). We have tested the model with 10,000

Table 42.4 Parameters for the ASAP platform

Model parameters	Description	Default	[Min,Max]
Population			
Population size	Number of adults at start of simulation	10,000	[1000,1000000]
Percentage of females	% of female agents	52	52
Maj/min split	% of individuals belonging to Maj	60	[60,90]
Neighborhoods	Number of neighborhoods	34	34
Death probability	Probability of dying each year	0.005	[0,0.1]
Age			
Min age Maj	Minimum number of living years for Maj	65	[55,75]
Max age Maj	Maximum number of living years for Maj	95	[70,100]
Mode age Maj	Mode value of life expectancy for Maj	80	[55–100]
Min age min	Minimum number of living years for min	45	[40,60]
Max age min	Maximum number of living years for min	85	[65–95]
Mode age min	Mode value of life expectancy for min	65	[40,95]
Education			
Max Edu Maj	Maximum education years for Maj	20	[15,25]
Min Edu Maj	Minimum education years for Maj	10	[8,15]
Max Edu min	Maximum education years for min	20	[15,25]
Min Edu min	Minimum education years for min	12	[8,15]
Marriage			
Min age marriage Maj	Minimum age for getting married Maj	26	26
Min age marriage min	Minimum age for getting married min	21	21
Marriage rate	Likelihood of getting married	0.02	[0.01,1]
Marriage age tolerance	Max age difference between agents	2	[1,5]
Marriage education tolerance	Max education difference between agents	2	[1,5]
Endogamy degree	Likelihood agent marries opposite group	0.05	[0,0.4]
Employment			
Number of employers	Number of job locations	5	[1,100]
Min friendly mode	Mode value of min friendly	0.85	[0,1]
Minority friendly	Degree of multicultural behaviors at job	Tri	[0,1]
Enforced antidiscrimination	Degree of antidiscrimination laws present at job	0.5	[0,1]

(continued)

Table 42.4 (continued)

Model parameters	Description	Default	[Min,Max]
Per Maj mal employed	% of males in Maj employed	1	[0,1]
Per Maj fem employed	% of females in Maj employed	0.7	[0,1]
Per min mal employed	% of males in min employed	0.85	[0,1]
Per min fem employed	% of females in min employed	0.35	[0,1]
Integration			
Suspicion Maj/min mean	Mean value of suspicion of Maj/min	0.5	[0,1]
Suspicion Maj/min SD	SD of suspicion of Maj/min value	0.25	[0,0.5]
Group support Maj/min mean	Mean value of in group support for Maj/min	0.5	[0,1]
Group support Maj/min SD	SD of in group support for Maj/min	0.25	[0,0.5]
Shared norms mean	Mean value of shared norms	0.5	[0,1]
Shared norms SD	SD of shared norms value	0.25	[0,0.5]
Personality traits			
HEXACO mean	Mean value for each HEXACO trait	0.5	[0,1]
HEXACO SD	SD of each HEXACO trait value	0.25	[0,0.5]
WorldView mean	Mean value of world view	0.5	[0,1]
WorldView SD	SD of world view value	0.25	[0,1]
Charisma mode	Agent's value of extraversion	Tri	[0,1]
Charisma min/max value	Agent's value of extraversion ±0.25	NA	[0,1]
Susceptibility	Agent's value of agreeableness	Tri	[0,1]
Susceptibility min/max values	Agent's value of agreeableness ±0.25	NA	[0,1]
Others			
Authority impact	Impact of authority level on social interactions	0.1	[0,1]
Interaction radius	Spatial radius within which individuals interact	500	[1,1000]

Maj majority, *Min* minority, *Tri* values from a triangular distribution (with Mode, Min, Max), *SD* standard deviation, *NA* not applicable.

agents interacting for 30 years. At 500 times the speed of a wall clock, it takes approximately 2.2 hours for a simulation to complete. We believe that we can enhance performance by improving the code, especially in the area of social network management.

42.7 Conclusion

Our research team's work on the artificial society analytics platform has been informed by our failures and successes during the process of developing several other models aimed at simulating societal dynamics [16–20]. We believe we are making progress. However, progress would be quicker if we could facilitate a debate on best practices in artificial society design. Such a debate could begin with the five design decisions presented in Sect. 42.2, but it need not end there. Every aspect of the construction process should be under consideration as we seek to provide better rationales for design decisions.

As the field of social simulation enters puberty, critical self-reflection involves "looking in the mirror" at our own artificial society design assumptions and inviting scrutiny from other experts. We have tried to foster this process by hosting sessions on "best practices" at conferences on modeling and simulation such as SpringSim 2019. This sort of self-reflection has the potential for inducing identity crises and for prompting moments of inspiration and insight. However, that's what will be required to guide our relatively new discipline safely through this transformative period, with its inevitable growing pains, until "social simulation" attains proper scientific adulthood.

Acknowledgments ASAP is being developed as part of the "Simulating Religion Project" at the Center for Mind and Culture in Boston and the "Modeling Religion in Norway" (MODRN) project, funded by The Research Council of Norway (grant #250449). Additional support has been provided by the Virginia Modeling, Analysis and Simulation Center in Suffolk, VA.

References

1. Conte, R., et al.: Manifesto of computational social science. Eur. Phys. J. Spec. Top. **214**(1), 325–346 (2012)
2. Squazzoni, F.: Agent-Based Computational Sociology, 2nd edn. Wiley, Hoboken (2012)
3. Wijermans, N., Jorna, R., Jager, W., van Vliet, T., Adang, O.: CROSS: modelling crowd behaviour with social-cognitive agents. J. Artif. Soc. Soc. Simul. **16**(4), 1 (2013)
4. Vanhee, F., Dignum, L.: Agent-based evolving societies. Adv. Intell. Syst. Comput. **229**, 283–294 (2014)
5. Xenitidou, M., Edmonds, B. (eds.): The Complexity of Social Norms, 2014. Springer, Cham (2014)
6. Elsenbroich, C., Verhagen, H.: The simplicity of complex agents: a contextual action framework for computational agents. Mind Soc. **15**(1), 131–143 (2016)
7. Flache, A., et al.: Models of social influence: towards the next frontiers. Jasss- J. Artif. Soc. Soc. Simul. **20**(4), 7 (2017)., issn: 1460–7425
8. Edmonds, B., Meyer, R.: Simulating Social Complexity: A Handbook. Springer International Publishing, Cham (2017)
9. Gilbert, N., Ahrweiler, P., Barbrook-Johnson, P., Narasimhan, K.P., Wilkinson, H.: Computational modelling of public policy. J. Artif. Soc. Soc. Simul. **21**(1), 1–19 (2018)

10. Voas, D., Bruce, S.: Secularization in Europe: an analysis of inter-generational religious change. In: Arts, W., Halman, L. (eds.) *Value Contrasts and Consensus in Present-Day Europe*. Brill, Leiden (2014)
11. Voas, D., Fleischmann, F.: Islam moves west: religious change in the first and second generations. Annu. Rev. Sociol. **38**, 525–545 (2012)
12. Aghababaei, N.: God, the good life, and HEXACO: the relations among religion, subjective well-being and personality. Ment. Health. Relig. Cult. **17**(3), 284–290 (Mar. 2014)
13. Gebauer, J.E., Bleidorn, W., Gosling, S.D., Rentfrow, P.J., Lamb, M.E., Potter, J.: Cross-cultural variations in big five relationships with religiosity: a sociocultural motives perspective. J. Pers. Soc. Psychol. **107**(6), 1064–1091 (2014)
14. Henrich, J.: The evolution of costly displays, cooperation and religion: credibility enhancing displays and their implications for cultural evolution. Evol. Hum. Behav. **30**(4), 244–260 (2009)
15. Lanman, J.A.: The importance of religious displays for belief acquisition and secularization. J. Contemp. Relig. **27**(1), 49–65 (2012)
16. Wildman, W.J., Sosis, R.: Stability of groups with costly beliefs and practices. J. Artif. Soc. Soc. Simul. **14**(3), 1–14 (2011)
17. Gore, R., Lemos, C., Shults, F.L., Wildman, W.J.: Forecasting changes in religiosity and existential security with an agent-based model. J. Artif. Soc. Soc. Simul. **21**, 1–31 (2018)
18. Shults, F.L., Lane, J.E., Diallo, S., Lynch, C., Wildman, W.J., Gore, R.: Modeling terror management theory: computer simulations of the impact of mortality salience on religiosity. Relig. Brain. Behav. **8**(1), 77–100 (2018)
19. Shults, F.L., Gore, R., Wildman, W.J., Lynch, C., Lane, J.E., Toft, M.: A generative model of the mutual escalation of anxiety between religious groups. J. Artif. Soc. Soc. Simul. **21**(4), 7 (2018). https://doi.org/10.18564/jasss.3840
20. Shults, F.L.: Can we predict and prevent religious radicalization? In: Øverland, G. (ed.) Violent Extremism in the 21st Century: International Perspectives, pp. 45–71. Cambridge Scholars Press, Cambridge (2018)

Chapter 43
Teaching the Complexity of Urban Systems with Participatory Social Simulation

Timo Szczepanska, Max Priebe, and Tobias Schröder

Abstract We describe how we use social simulation as a core method in a new master's program designed to teach future leaders of urban change to deal with the complexity inherent in current societal transformations. We start by depicting the challenges with regard to cross-disciplinary knowledge integration and overcoming value-based, rigid thinking styles that inevitably arise in the process of solving ecological, technological, or social problems in cities new and old. Next, we describe a course based on urban modeling and participatory approaches, designed to meet those challenges. We reflect on our first experience with this approach and discuss future developments and research needs.

Keywords Urbanization · Complexity · Systems thinking · Higher education · Participatory modeling

43.1 Introduction

For a few decades, academics and policymakers have emphasized the importance of understanding the complex functioning of urban settlements [1–4]. While much research has been devoted to (re-)formulating social, regulatory, or environmental challenges and opportunities of the asserted urban age in respect to particular fields of interest [5], cross-disciplinary knowledge integration remains an ongoing challenge.

In this article, we want to focus on how the community of social simulation scholars can help to equip the next generation of experts in the field of urban planning and administration with comprehensive tools that foster a systems-oriented way of understanding, in order to adequately meet the challenges that arise from an ever-increasing complexity of urban systems. To this end, we propose a course

T. Szczepanska (✉) · M. Priebe · T. Schröder
University of Applied Sciences Potsdam, Potsdam, Germany
e-mail: timo.szczepanska@uit.no

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_43

outline designed (and tested twice so far) for first-term students in a novel inter- and transdisciplinary master's program called Urban Futures at the Potsdam University of Applied Sciences in Germany [6]. One particular challenge lies in the fact that the majority of students in said program have undergraduate degrees in the arts and humanities, mostly lacking a quantitative mindset (despite enthusiasm to learn new methods); only very few of them have interacted with computer code before.

Our main aim of approaching matters of urban complexity together with a multidisciplinary group of students by means of a modeling course is to make different assumptions transparent, replicable, and thus discussable. Here, modeling is understood as a process of abstraction and construction [7], helping to formulate statements about properties, relations, and contexts of social phenomena in urban settings. We follow Epstein's claim [8] that any given intellectual statement is underpinned by implicit models, which can only be fully understood by others – and, ultimately, oneself – if made explicit. Explicitness in this context serves two goals. First, it encourages interdisciplinary communication between students from various backgrounds fostering cross-disciplinary knowledge integration. Second, explicitness forces students to distinguish clearly between their conceptual understandings and their normative assertions. We expect that the latter practice can ultimately lead the way from an individualistic and normative perspective upon urban topics to a commonly shared understanding that helps to coordinate actions and prevent conflicts. Thus, our hope is to use social simulation as a tool to instill in students a scientific mindset assisting them to overcome the shortcomings of ideological, closed-minded thinking styles that impede the creation of accurate knowledge about the future [9].

In the following section, we provide an overview of the concept and objectives underpinning the Urban Futures program and our aims in approaching urban complexity through formal modeling and social simulation. Subsequently, Sect. 43.3 provides insights into our rationale to selecting a particular modeling approach and further depicts the proposed course outline in detail. It is followed by concluding remarks and an outlook for future developments and research needs.

43.2 Urban Complexity in Higher Education: Knowledge Integration and Systems Thinking

The twenty-first century is the century of cities. For the first time in the history of humanity, the majority of our species live in urban environments as opposed to the small-scale, rural communities that have been most humans' habitat since the agricultural revolution [1]. Urbanization is an ambiguous phenomenon. On the one hand, the close connectedness and related exchange of knowledge and information that are possible in cities have brought unprecedented wealth, prosperity, and solutions to the most devastating problems humanity has been plagued with for its entire existence such as famine, extreme violence, and infectious disease [10].

On the other hand, the human mind, evolved to be adaptive in a natural, small-group environment, is not well equipped to deal with the complexity inherent in the urban way of life [11]. The brain's limited information-processing capacities, its preference for coherent narratives and confirmatory thinking, as well as overreliance on social metaphors and resulting tribal tendencies get in the way of the rational, scientific ways of thought required for sound decision-making about the future of a complex system such as a city [9, 10]. However, we cannot train every stakeholder to be a scientist at the PhD level, and we cannot wait with many decisions that have long-lasting consequences until they have been studied scientifically from each and every angle.

With this problem in mind, the Urban Futures master's program at the Potsdam University of Applied Sciences was developed to train a future generation of change agents (or transformation managers) to approach the "wicked problems" [12], inherent in twenty-first century urban transformations with a scientific approach that is highly adaptable to context and scale and aimed at methodologically stringent knowledge integration, both across relevant scientific disciplines and between academics and practitioners. Building onto ideas about and experiences with "transformatory science" [13], "research-based learning" [14], and "transdisciplinary case studies" [15] popular in the German higher-education context, the Urban Futures curriculum is centered around project-based courses. Students work in small, interdisciplinary groups on mostly self-chosen practical problems related to sustainable development of cities. Professors provide guidance and support, but (mostly) refrain from traditional lecture formats. A component of the curriculum aimed at preparing and supporting the core project-based activities consists of a series of smaller courses devoted to specific methods that are thought to be useful not only for the self-organized research projects but also for similar challenges in students' later professional lives. One of these courses is a modeling course aimed at using spatially explicit social simulation to foster scientific thinking and better inter- and transdisciplinary communication about the city as a complex system.

43.3 Integrating Urban Modeling in the Curriculum

43.3.1 Rationale for the Course

We describe here a course design for a 3 ECTS point course to be offered in the upcoming Winter term 2018–2019. The course design builds on the experience gained with a preliminary version of the course offered in the previous 2 years of our master's program since its inception in 2016. One major issue for students identified by evaluating the course consisted in the high level of abstraction of addressed models, which resulted in some students perceiving modeling a methodological obscurity rather than a general-purpose tool for understanding. Consequently, our aim for the redesign of the course is to increase the explanatory value of urban modeling approaches by emphasizing their field of application and contextualizing models with real-world examples from the Berlin–Brandenburg area surrounding the university.

In preparation for this course, we scouted the landscape of different modeling approaches to urban systems [16]. Our findings suggest that, similar to the realm of landscape ecology, one fundamental value of urban modeling consists in its spatial explicitness. The spatial dimension of urban models enables modelers to contextualize their work and adequately mimic selected spatial features – both of locations and relations – of an observed, real system. For urban models, simulating aspects of autopoietic complex system that are per se constituted by features distributed in space, spatial explicitness is thus crucial. Additionally, scholars have emphasized that cities need to be understood not only by their spatial properties but also as entities pursuing dynamic change in a state of constant disequilibrium [17], thus calling for an integrated spatial–temporal perspective.

Typically, the modeling concept of cellular automata (CA) is used for studying these complex dynamics achieving the incorporation of spatial and temporal aspects at the same time. Urban CA is a discrete model concept consisting of a regular grid of cells, each of which is capable of expressing different states. It has originated at the intersection of computer science (John von Neumann's theory of self-replicating automata) and geographical science (Tobler's cellular models) since the 1960s [18]. Prominent use cases of urban CA can be found for subjects such as land use change [19, 20], urban growth [21], and urban sprawl [22, 23]. Despite its popularity in the field of urban modeling, developments in state-of-the-art research suggest a move away from classic CA approaches – with their at-times disadvantageous reliance on fixed neighborhoods and relatively simplistic decision-making algorithms – toward a spatially explicit, agent-based paradigm for urban models [24]. This shift is being supported by a steady increase of computational performance, newly developed software environments, the increase in availability of geographical information data, and the growing prominence of agent-based models (ABM) as a general framework for social simulation.

Our particular rationale behind choosing spatially explicit models for integration in the curriculum is motivated by calls from students and partners to provide tools that are suitable for application in real-world contexts as opposed to conceptual tools designed for debates in academia. Moreover, we consider spatially explicit models as easier to be understood by stakeholders – a hypothesis we want to test empirically in the near future by means of supervised student research in the field. Thus, one major change for the course is a switch from NetLogo to the GAMA Platform [25] as the primary modeling tool and our ability to rely on existing models of real cities in the Berlin–Potsdam area, which we have started to develop and expect to grow in number over time.

The course consists of five working sessions with a duration of 3 hours each (Fig. 43.1). Additionally, there are two presentation sessions scheduled and the requirement to develop an individual, written model concept. The purpose of the course is to join an understanding of ABM logic and learn how to conceptualize models. At the end of the course, the students of the master's degree program are

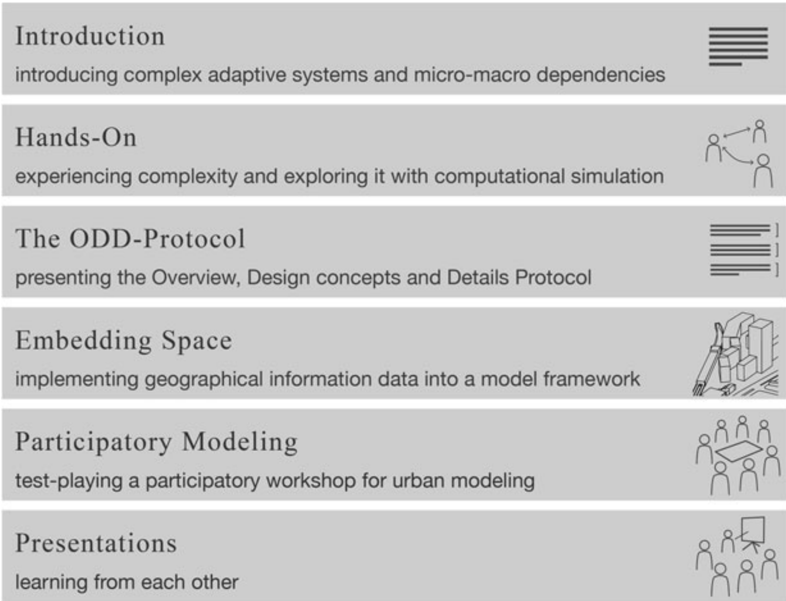


Fig. 43.1 Roadmap of the course outline

supposed to have gained a comprehensive understanding of urban phenomena as complex adaptive systems [26]. In addition, they will develop the concept for an own model idea in the form of a standardized protocol [27]. Students do not learn to program in the course, but they can choose an optional introductory course on programming (for data sciences) in the following semester. The logical structure of the course conveys the different design concepts of an agent-based simulation, where each session focuses on one specific context that introduces central elements of urban modeling.

43.3.2 Introducing Complex Systems and Agent-Based Modeling

The first session is used to introduce the idea of complex adaptive systems and to show how urban phenomena can be explored by analyzing micro–macro dependencies. It is subdivided into three sections where the class will approach the broader topic of agent-based modeling step by step.

1. Section one addresses the general question on why one should model. As an icebreaker, we read out loud a paragraph of Michael Ende's book *Momo* (a widely known children's book in Germany) to trigger an open discussion on the purpose of modeling: *The tyrant Marxentius Communus "the Red" was*

disappointed with the whole world of that time. He orders his workers to build a new one that should be as big as the old one. When the workers started to use the material from the old globe to build the new one not only the new globe grew but also the old earth became smaller. One day a worker moved the last pebble to his new location, and with this, the old world disappeared. Next, we move from fiction to science. On the basis of the short paper "Why should we model?" by Joshua Epstein [8], we discuss the "generative method" and the idea to use models as a communication tool to make implicit knowledge explicit.

2. Section two deals with questions about the nature of ABM [28, 29] by introducing the idea of emergence and the macro–micro paradigm [29, 30]. To prepare students to explore agent-based models by themselves, we elaborate on basic features such as the bottom-up approach to generate social phenomena. At the end of the section, they have learned how an agent-based model controls a set number of software agents with predefined rules of conduct that interact on a local level and how their predefined interaction capacities (micro) generate emerging phenomena (macro).
3. Section three is dedicated to the question of how one can explore the complex relationships between physical, socialcultural, and informational fields of action of urban systems. This question will be tackled using the example of two model types that cope with social phenomena in urban spaces. First, Thomas Schelling's [31] segregation model, a readily comprehensible checkerboard model, is used to illustrate how microlevel choices of individual agents can lead to the formation of unintended emergent macro-level patterns. A surprising result of this model of segregation is that even a slight preference in choosing a place of residence—namely, the preference of an agent not to belong to a minority in his immediate neighborhood—is sufficient to generate a distinctly segregated model world [4]. Second, referring to the stylized framework of an urban model (see Fig. 43.2), the class explores the differences between agent, interaction, and environment.

43.3.3 Hands On

The goal of the third session is the introduction and exploration of the GAMA-Platform by a combination of playful performance and a computational analysis. The class starts with the installation of the GAMA-Platform and a brief introduction of key features of the user interface. After a short break, we continue with a warm-up game of "Friends and Foes" which was introduced by the Fratelli Group at the Cambridge conference "Embracing Complexity" in 1999. The game is easily adaptable for classroom settings and a fun way for students to make own experiences with the emergence as different behaviors will directly result in different patterns shown by the entire group. The game was originally implemented as "Heroes and Cowards" [30] in NetLogo, which we reproduced for the GAMA-Platform. The combination of a playful approach and the computer model enables the class to

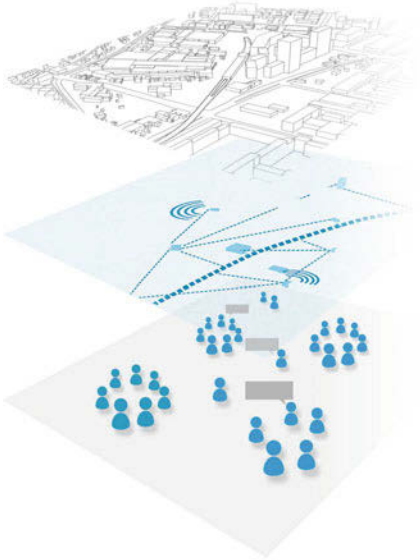


Fig. 43.2 Stylized framework of an urban model

research their real-world observation in an abstracted model. To start the game, we instruct the students to prepare a squared restricted playing field in the classroom. In the beginning, every student has to choose two random classmates and keep their choice secret until the end of the game. The first choice represents the personal antagonist, the second the personal friend. Our version of the game consists of three consecutive rounds with a diverse set of rules. Every round starts with the students spreading out randomly throughout the room.

In the first round, everyone is asked to behave timidly. The only obligation one has is to always position oneself in between the friend and the enemy. As soon as a stable pattern has emerged or looping situation arises, we gather and reflect on the students' observations. In the second round, everybody is asked to behave bravely by positioning oneself between the friend and the enemy. Before the second round starts, we ask the students what they expect to happen with the new ruleset. After a stable or repeating situation occurs, we stop the game and compare expectations with the actual observations. In the last round, we divide the students into two equal groups. Half of them behave bravely, the other half timidly. Again, we ask the students what they expect to happen this time and start the new round to gather observations. Finally, we discuss what happened and if assumptions at the beginning of each round matched the observation.

Now we turn to the digital version. We open the corresponding "Heroes and Cowards" model in the GAMA application and repeat the three experiments in the model. After some playing time, the class comes back to the questions and assumptions gathered before and discusses them in the plenum. For the next session, the students read the article “The dissemination of culture, a model with local convergence and global polarization” [32].

43.3.4 The ODD Protocol

In this session, we introduce students to the standardized *Overview, Design Concepts and Details Protocol*, short ODD Protocol [27, 33, 34]. The primary goal of the session is to work out how to describe model concepts in the form of an ODD Protocol and to reveal the underlying thoughts in a comprehensible way. When working with the ODD Protocol, the components and concepts of an ABM already discussed in the class can be reviewed alongside the subitems of the protocol. The revisiting helps the students to reiterate the previously learned in a structured and comprehensive manner. By learning to create model concepts in the framework of the ODD, we make sure that our students reflect explicitly on the elements of their model ideas and relevant modeling decisions, allowing them to discuss their models in the classroom on basis of a shared understanding.

We start with a presentation on ODD Protocol with provided model examples. In the following, each student is asked to individually work on an ODD draft based on their understanding of the article “The dissemination of culture: A model with local convergence and global polarization” [32]. A template for the ODD Protocol is provided. In the last section of this session, the students present their draft to a partner. Finally, we discuss in the plenum what information can be easily extracted from the text and what is missing to complete the ODD. For the next session, each student gets the task to formulate first thoughts for an own model concept.

43.3.5 Embedding Space

The objective of this session is to implement geographical information data into a model framework to ensure the integration of spatial dimensions within the modeling course. The shown model of a partner city (Fig. 43.3) is an example to explain how to transform a selected area from the OpenStreetMap interface (www.openstreetmap.org) into a proper model environment. Topics of this course component comprise an introduction to OpenStreetMap (OSM) including an explanation of the basic structure of OSM files, essential attributes, and how to address them. The session ends with a hands-on lecture on how to select an area and implement it into GAMA.

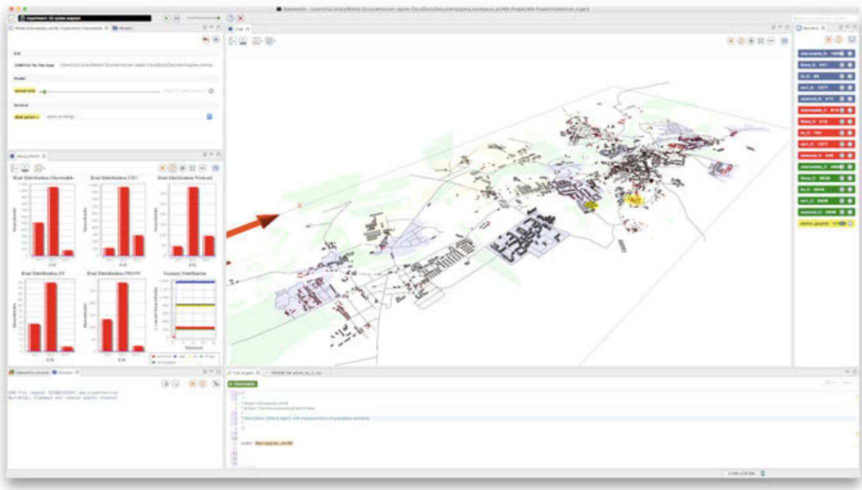


Fig. 43.3 Embedding OSM Data into GAMA underlined with population statistics

43.3.6 Participatory Modeling

The aim of this course component is to cooperatively test-play the design of a participatory workshop for urban modeling. It commences with a brief input on the objectives of the existing participatory modeling approaches, drawing from group modeling within organizational studies [35] and selected cases of the companion modeling approach [36]. Students are given examples that illustrate different steps, roles, and problems typically involved in participatory modeling workshops. After a short Q&A session, students are confronted with a campus-related problem. They are encouraged to collectively develop a solid and shared foundation of a prototypical domain model. In order to kick-off this "discovery phase" [37], students are asked to individually identify relevant system elements, share their thoughts with their neighbors, and finally present them on paper to the whole class. In the following hour, students are asked to first cluster the identified elements, second locate the different elements with their corresponding variables on a map (Fig. 43.4), and third conceptualize agent types together with their decision-making principles. During the last hour, a visualization table (Fig. 43.4) and a generic GAMA model of the campus are employed in order to explore and discuss the obtained results of the prototyping exercise. User–model interaction via the user interface enables the students to playtest their assumptions and discuss the visualized output without having to manipulate the code. This process is loosely inspired by serious gaming approaches that facilitate experimenting and learning in novel ways [38, 39].

The collaborative exercise is supposed to trigger motivation and inspiration for the homework, that consists of writing an ODD for a GAMA model. At the end of this lesson, students get the mandatory assignment to develop their first rough

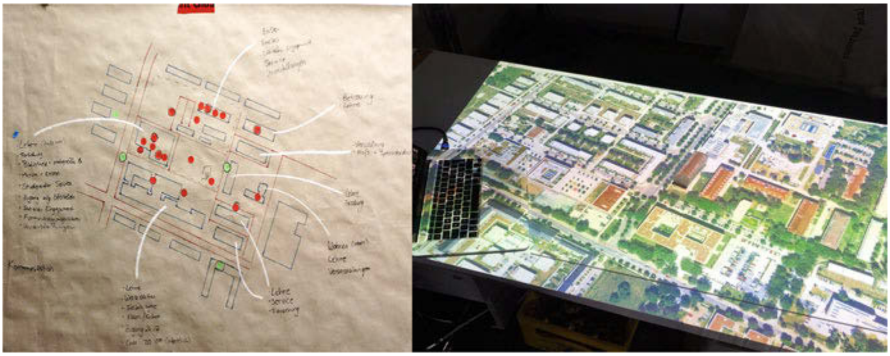


Fig. 43.4 Locating elements on a map and play-test the model

model idea into a model concept and present it to the course. For this presentation, 20-minute units will be scheduled for the next two sessions.

43.3.7 Presentations of Model Ideas

The last two sessions are designated for the students' presentations of their model concepts. During the talks, students can benefit from feedback of both the lecturer and classmates in the audience. The sessions will be concluded by a Q&A section where the students can address difficulties that arose during the development process and latest issues concerning the paragraphs of the ODD Protocol. With the gathered feedback, students are prepared to finalize their model idea in the form of an ODD Protocol. Experience with the two preliminary versions of the course carried out so far shows that the feedback sessions and discussion stimulated through the presented model concepts are especially valuable for attaining the overarching educational goal of fostering model-based systems thinking when approaching urban problems.

43.4 Conclusions

With the course presented here, we hope to provide students with an insight into state-of-the-art social simulation tools for urban modeling. Moreover, we try to encourage students to apply these tools to contexts they are familiar with and problems they consider to be meaningful. As we have argued in this chapter, we see the need to prepare our students to adequately understand urban dynamics in a matter that promotes scientific reasoning and allows for cross-disciplinary knowledge integration. By training these skills with cutting-edge participatory urban-modeling technology, we hope to increase our students' ability to anticipate, successfully respond to, and intentionally shape urban transformations in the future. The urban modeling course outline presented here is just one of the many pillars in this venture. Additionally, offered courses cover topics such as data visualization, forecasting and methods, and transformation management. Following the method-oriented part of the curriculum, a field study project allows students to apply their learnings in a workshop-like setting at one of the university's partner cities. We regard these field studies not only as an opportunity for participants of the master program but also as a part of our mission as a University of Applied Sciences to foster the regional diffusion of knowledge and innovation, providing mid-sized and smaller municipalities with services and tools that are too often reserved to bigger and wealthier cities pursuing the objectives of smart-city initiatives.

In order to provide empirical evidence for our claim that the collaborative process of participatory modeling as mentioned in this chapter can benefit both higher education and professional practice in the sphere of urban planning and administration, we are preparing to accompany our activities from 2019 onwards with rigorous, social science-oriented impact assessments. Thus, future research will allow us to identify potential pitfalls and iteratively integrate new learnings in the participatory modeling process. Until then, we are looking forward to receiving valuable feedback from students and the scientific community.

Acknowledgments We acknowledge funding from the European Funds for Regional Development, grant number 85009319.

References

1. UN-Habitat (United Nations Human Settlement Programme). World Cities Report 2016. Urbanization and Development: Emerging Futures (2016). https://unhabitat.org/books/world-cities-report/
2. Batty, M.: The New Science of Cities. MIT Press, Cambridge, MA (2013)
3. Wilson, A.G.: Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis. Pearson Education, Harlow (2000)
4. Crooks, A.T., Malleson, N., Wise, S., Heppenstall, A.: Big data, agents and the city. In: Schintler, L.A., Chen, Z. (eds.) Big Data for Urban and Regional Science, pp. 204–213. Routledge, New York (2018)
5. Burdett, R., Sudjic, D.: The Endless City. Phaidon, London (2006)
6. Prytula, M., Schröder, T., Dörk, M., Ortgiese, M., Michel, A., Neuroth, H., Lehmann, J.: Transformation lernen: Der Masterstudiengang Urbane Zukunft an der Fachhochschule Potsdam. Transform. Cities. **3**, 76–79 (2016)
7. Ihde, D.: Models, Models Everywhere. In: Lenhard, J., Küppers, G., Shinn, T. (eds.) Simulation. Sociology of the Sciences Yearbook, vol. 2, pp. 79–86. Springer, Dordrecht (2006). https://doi.org/10.1007/1-4020-5375-4_5
8. Epstein, J.M.: Why model? J. Artif. Soc. Soc. Simul. **11**(4), 12 (2008). http://jasss.soc.surrey.ac.uk/11/4/12.html
9. Tetlock, P.E., Gardner, D.: Superforecasting: The Art and Science of Prediction. Random House, London (2015)
10. Pinker, S.: Enlightenment Now: The Case for Reason, Science, Humanism, and Progress. Penguin, London (2018)

11. Miller, G.: Roots of the urban mind. Science. **352**(6288), 908–911 (2016). https://doi.org/10.1126/science.352.6288.908
12. Rittel, H.W.J., Webber, M.M.: Dilemmas in a general theory of planning. Policy. Sci. **4**, 155–169 (1973). https://doi.org/10.1007/BF01405730
13. Schneidewind, U.: Transformative Wissenschaft-Motor für gute Wissenschaft und lebendige Demokratie. GAIA-Ecol. Perspect. Sci. Soc. **24**(2), 88–91 (2015). https://doi.org/10.14512/gaia.24.2.5
14. Mieg, H.A., Lehmann, J.: Forschendes Lernen: Wie die Lehre in Universität und Fachhochschule erneuert werden kann. Campus-Verlag, Frankfurt (Main) (2017)
15. Scholz, R.W., Lang, D.J., Wiek, A., Walter, A.I., Stauffacher, M.: Transdisciplinary case studies as a means of sustainability learning: historical framework and theory. Int. J. Sustain. High. Educ. **7**(3), 226–251 (2006). https://doi.org/10.1108/14676370610677829
16. Priebe, M., Szczepanska, T., Schröder, T., Prytula, M., Dörk, M.: Potentiale partizipativer Systemmodellierung in Brandenburger Kommunen. IaF Working Paper (2017). https://www.fh-potsdam.de/fileadmin/user_dateien/5_forschen/Publikationen/201711_Arbeitspapier_Pasymo.pdf
17. Batty, M.: Cities and Complexity: Understanding Cities with Cellular Automata, Agent-Based Models, and Fractals. MIT Press, Cambridge (2005)
18. Iltanen, S.: Cellular automata in urban spatial modeling. In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (eds.) Agent-Based Models of Geographical Systems, pp. 69–84. Springer, Dordrecht (2012). https://doi.org/10.1007/978-90-481-8927-4_4
19. Batty, M., Xie, Y., Sun, Z.: Modeling urban dynamics through GIS-based cellular automata. Comput. Environ. Urban. Syst. **233**, 205–233 (1999). https://doi.org/10.1016/S0198-9715(99)00015-0
20. White, R., Engelen, G.: Cellular-automata and fractal urban form - a cellular modeling approach to the evolution of urban land-use patterns. Environ. Plan. A. **25**(8), 1175–1199 (1993). https://doi.org/10.1068/a251175
21. Benenson, I., Torrens, P.M.: Geosimulation: Automata-Based Modeling of Urban Phenomena. Wiley, London (2004). https://doi.org/10.1002/0470020997
22. Batty, M., Crooks, A.T., See, L.M., Heppenstall, A.J.: Perspectives on agent-based models and geographical systems. In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (eds.) Agent-Based Models of Geographical Systems, pp. 1–15. Springer, Dordrecht (2012). https://doi.org/10.1007/978-90-481-8927-4
23. Felsen, M., Wilensky, U.: NetLogo Urban Suite – Sprawl Effect Model. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston (2007). http://ccl.northwestern.edu/netlogo/models/UrbanSuite-SprawlEffect
24. O'Sullivan, D., Millington, J., Perry, G., Wainwright, J.: Agent-based models – because they're worth it? In: Heppenstall, A., Crooks, A., See, L., Batty, M. (eds.) Agent-Based Models of Geographical Systems, pp. 109–123. Springer, Dordrecht (2012). https://doi.org/10.1007/978-90-481-8927-4_6
25. Taillandier, P., Vo, D.A., Amouroux, E., Drogoul, A.: GAMA: a simulation platform that integrates geographical information data, agent-based modeling and multi-scale control. In: Desai, N., Liu, A., Winikoff, M. (eds.) Principles and Practice of Multi-Agent Systems. PRIMA 2010. Lecture Notes in Computer Science, vol. 7057, pp. 242–258. Springer, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-25920-3_17
26. Miller, J.H., Page, S.E.: Complex Adaptive Systems: An Introduction to Computational Models of Social Life. Princeton University Press, Princeton (2007)
27. Grimm, V., Polhill, G., Touza, J.: Documenting social simulation models: the ODD protocol as a standard. In: Edmonds, B., Meyer, R. (eds.) Simulating Social Complexity. Understanding Complex Systems, pp. 117–133. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-540-93813-2_7
28. Edmonds, B., Meyer, R.: Simulating Social Complexity: A Handbook. Springer, Berlin, Heidelberg (2013). https://doi.org/10.1007/978-3-540-93813-2

29. Crooks, A.T., Heppenstall, A.J.: Introduction to agent-based modeling. In: Heppenstall, A.J., Crooks, A.T., See, L.M., Batty, M. (eds.) Agent-Based Models of Geographical Systems, pp. 85–105. Springer, Dordrecht (2012). https://doi.org/10.1007/978-90-481-8927-4_5
30. Wilensky, U., Rand, W.: An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo. MIT Press, Cambridge (2015)
31. Schelling, T.: Micromotives and Macrobehavior. W. W. Norton, London & New York (1978)
32. Axelrod, R.: The dissemination of culture, a model with local convergence and global polarization. J. Confl. Resolut. **41**, 203–226 (1997). https://doi.org/10.1177/0022002797041002001
33. Grimm, V., Berger, U., Bastiansen, F., Eliassen, S., Ginot, V., Giske, J., Goss-Custard, J., Grand, T., Heinz, S., Huse, G., Huth, A., Jepsen, J.U., Jørgensen, C., Mooij, W.M., Müller, B., Peer, G., Piou, C., Railsback, S.F., Robbins, A.M., Robbins, M.M., Rossmanith, E., Rüger, N., Strand, E., Souissi, S., Stillman, R.A., Vabø, R., Visser, U., DeAngelis, D.L.: A standard protocol for describing individual-based and agent-based models. Ecol. Model. **198**, 115–126 (2006). https://doi.org/10.1016/j.ecolmodel.2006.04.023
34. Grimm, V., Berger, U., DeAngelis, D.L., Polhill, G., Giske, J., Railsback, S.F.: The ODD protocol: a review and first update. Ecol. Model. **221**, 2760–2768 (2010). https://doi.org/10.1016/j.ecolmodel.2010.08.019
35. Vennix, J.: Group Model Building: Facilitating Team Learning Using System Dynamics. Wiley, New York (1996)
36. Etienne, M.: Companion Modelling: A Participatory Approach to Support Sustainable Development. Springer, Dordrecht (2014). https://doi.org/10.1007/978-94-017-8557-0
37. Williams, R.A.: Lessons learned on development and application of agent-based models of complex dynamical systems. Simul. Model. Pract. Theory. **83**, 201–212 (2017). https://doi.org/10.1016/j.simpat.2017.11.001
38. Verhagen, H., Johansson, M., Jager, W.: Games and online research methods. In: Fielding, N., Lee, R., Blank, G. (eds.) The SAGE Handbook of Online Research Methods, 2nd edn. SAGE Publications Inc., London (2017). https://doi.org/10.4135/9781473957992.n17
39. Gugerell, K., Zuidema, C.: Gaming for the energy transition. Experimenting and learning in co-designing a serious game prototype. J. Clean. Prod. **169**, 105–116 (2017). https://doi.org/10.1016/j.jclepro.2017.04.142

Chapter 44
Enabling Innovation Within Public Research Institutes: A Modelling Approach

SSC2018 – Post-proceedings (April 2019)

Davy van Doren

Abstract Over time, the output produced by public research institutes (PRIs) has become increasingly scrutinised. As PRIs are expected to facilitate the realisation of innovation policy and align with the grand challenges of our time, they constantly need to assess their effectiveness. In this chapter, we present a conceptual framework on how the management of organisational processes and capabilities that enable innovation can be modelled through combining a system dynamics and agent-based modelling approach. Focusing on collaboration mechanisms with potential target groups, we describe how this conceptual framework has been formalised in simulation based on different empirically observed PRI types and strategies. Our model could provide PRIs with valuable insights into their role within collaboration structures and indicate how the operation of different PRIs could be coordinated to optimise the output of innovation systems.

Keywords Agent-based modelling · System dynamics · Public research · Organisational strategy · Innovation

44.1 Introduction and Theoretical Background

With innovation often being a central pillar within public policy for driving development, there is increased pressure for public research institutes (PRIs) to demonstrate how they contribute [1]. As such, there is an evolving need how PRIs can assess their R&D activity and find ways how their roles and effectiveness within technological trajectories can be optimised. However, due to the diversity of existing

D. van Doren (✉)
IQIB Institut für qualifizierende Innovationsforschung & -beratung, Bad Neuenahr-Ahrweiler, Germany
e-mail: Davy.vanDoren@iqib.de

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_44

PRIs – for example, regarding organisational objectives, available resources or embedded capabilities – there is unlikely to be a one-size-fits-all solution. As such, to improve understanding on the functioning of PRIs as enablers of innovation, it is necessary to identify and analyse the effects these institutes have on the dynamic of innovation externally and how these effects relate to their internal operation.

One process to be considered when analysing the innovation performance of PRI regards how collaboration with potential clients or target groups is being managed [2–6]. On the one hand, PRIs are organisations that have explicitly and consciously created purposes; that are being managed based on resource generation, allocation and optimisation; that consist of formal competing structures and that are functioning through processes influenced by strategies, routines, standards and norms. Leaders and managers are generally well-informed regarding the organisational capabilities and operations and can control the allocation of available resources to achieve set goals. On the other hand, PRIs are also strongly dependent on the needs and demands from various external stakeholders. As PRIs are often considered as being intermediaries for innovation – through responding to external R&D initiatives and acting based on impulses provided by target groups – PRIs need to cope with multiple actors within an environment characterised by uncertainty. Due to bounded rationality concerning an evolving environment resulting from stochastic and often uncoordinated activities, PRIs need to develop capacity on how the often uncertain and changing external demands can best be managed and aligned with internal operation [7, 8].

To better understand the role that individual PRIs can play in enabling innovation and how internal operation could be aligned with the activity of other stakeholders to optimise the overall output of innovation systems, it is likely that PRIs require the capability to identify and address changing needs of target groups. As such, we argue that PRIs that invest resources to optimise collaboration with target groups will increase both individual fitness and the systematic output of innovation.

Hypothesis: PRIs that make explicit investments to optimise collaboration with target groups increase both individual fitness and the systematic output of innovation.

44.2 Methodological Approach

Research is conducted within the context of the project ‘Enabling Innovation by Simulation’ (EIS).[1] EIS builds conceptually on the tool ‘Enabling Innovation’, which is a management methodology previously developed for PRIs to improve the identification and improvement of organisational capabilities and processes in the

[1] https://www.iqib.de/de/eis-enabling-innovation-by-simulation-simulation-der-innovationsfahigkeit-von-instituten.html

context of innovation [9]. To gain understanding how the performance of PRIs can be assessed on a more formal basis, the development of a new simulation model was initiated. Simulation models can be valuable to obtain more understanding regarding the action and interaction of internal and external processes. Also, simulation models can provide a space for experiments and can operate on wide spatial-temporal scales [10, 11].

The developed EIS model contains both a system dynamics (SD) and an agent-based model (ABM) dimension [12–14]. The SD dimension represents the *metabolism* of PRIs. The SD dimension implies that existing relationships, processes and resources can be rationally coordinated or allocated within organisations. Learning and strategy-making processes can either be executed or updated based on internal decision-making criteria or existing organisational management regimes. The ABM dimension represents the environment – or *innovation system* – in which PRIs operate [15]. With behavioural rules representing the formal internal logic of PRIs to (inter)act within its environment, it also represents their limited ability to predict or control the dynamic of innovation systems. The ABM dimension also constitutes the stochastic and spatiotemporal *movement* and *distribution* of agents.

The model has been mainly developed in a deductive fashion – that is, the use of theoretical assumptions and empirical data in building the logic of a model. Quantitative data was collected and analysed for identifying relationships between different actors and for statistically identifying and distinguishing between actor types. Quantitative data included publications, patents, research projects and information on organisation size, budget allocation and spin-off activity. Qualitative data was collected by means of interviews with representatives of different PRIs. Qualitative data was used for obtaining contextual insights regarding the behaviour of PRIs, particular regarding the nature and extent of existing organisational processes, capabilities and strategies used in the context of innovation. Interviews were transcribed, coded and analysed [16].

Figure 44.1 depicts the ABM dimension representing the innovation system in which PRIs operate, whereas Fig. 44.2 conceptualises the SD dimension representing the organisational metabolism of PRIs. The software platform Netlogo[2] was used for model development.

The developed simulation model was applied to assess the impact of different PRI collaboration strategies. Instead of performing extensive model calibration steps, the model has been mainly applied in an inductive fashion – that is, the interpretation of generated in silico data for understanding the performance of different strategies. Strategies were defined using insights derived from the analysis of various distinguished PRI types. Selected strategies were then tested by means of conducted simulation experiments.

[2]https://ccl.northwestern.edu/netlogo/

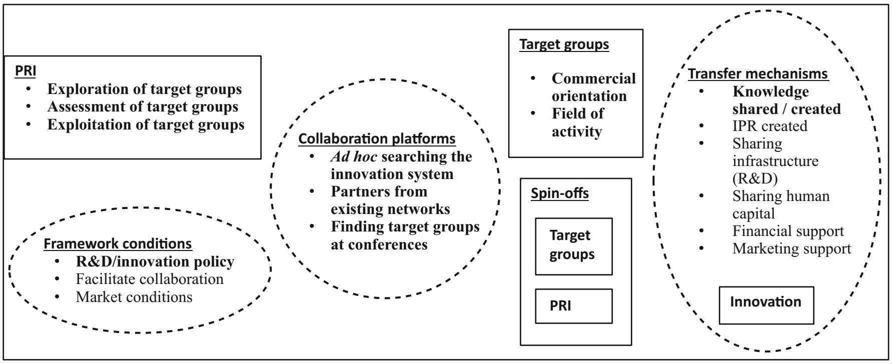


Fig. 44.1 Conceptual model for PRI external processes – innovation system (ABM). Within the modelled innovation system, there are two types of agents: *PRIs* and *target groups*. Both PRIs and target groups strive towards producing innovations. In order to produce innovations, agents require a certain amount of knowledge, of which there are two types: (1) *fundamental knowledge* and (2) *applied knowledge*. There are various ways how agents can obtain knowledge: (1) agents can obtain fundamental knowledge by means of investing into internal learning processes, (2) agents can obtain fundamental knowledge through exploring the environment (innovation system) or (3) agents can obtain knowledge through collaboration with other agents. To collaborate with target groups, PRIs need to carry out three procedures to establish a successful collaboration: (1) find/explore potential partners, (2) assess potential partners and (3) collaborate/exploit with partners. For PRIs, successful collaborations are the only means to obtain applied knowledge. There are three modes how collaborations can be established by PRIs: (1) PRIs can simply explore the environment searching for potential partners, (2) PRIs can find partners through participating in conferences and (3) PRIs can establish collaborations through networks, which contain previous collaboration partners. The squared components have agency within the ABM. The model elements that have been translated into the EIS simulation model are in bold

44.3 Results

The following are the results of three selected investigated strategies:

- *Strategy 1 – Invest in internal dedicated organisational processes:* A PRI's implementation of exploration, assessment and exploitation processes to improve collaborations with target groups.
- *Strategy 2 – Participate in conferences:* A PRI's participation in conferences, considering both the price of participating in conferences and the consistency/frequency of participation in conferences.
- *Strategy 3 – Participate in networks:* A PRI's participation in networks, considering both the price of participating in networks and the consistency/frequency of participation in networks.

Overall, the following observations were made:

Invest in internal dedicated organisational processes: When no investments are made in any of the four investigated collaboration parameters, PRIs perform weakly. PRI fitness, the amount of collaborations and the amount of created

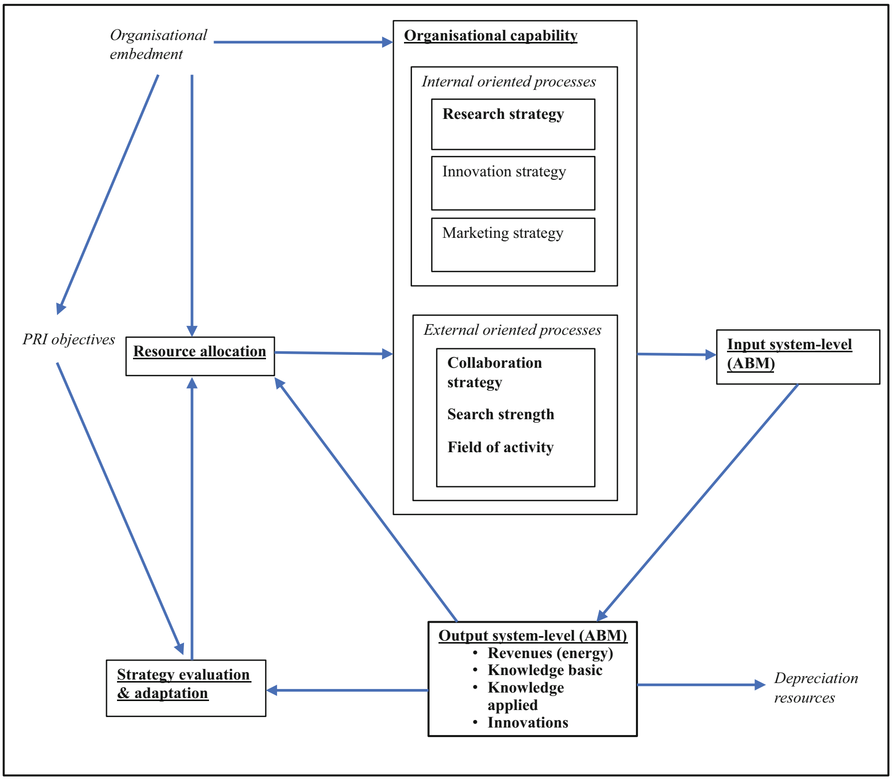


Fig. 44.2 Conceptual model for PRI internal processes – metabolism (SD). To optimise collaboration processes, PRIs can invest resources to improve certain capabilities required for successful collaboration. As capabilities are subject to degeneration, they require investments to remain functional or to improve. Four capability variables are distinguished: (1) exploration strength: the ability to move around and explore the environment, (2) collaboration strength: The ability to initiate a collaboration once a partner has been found, (3) assessment strength: the ability to assess the collaboration match between a PRI and a found partner and (4) exploitation strength: the ability to maximise gains obtained from a collaboration. PRIs can follow different strategies how investments are made in the capability variables. Also, PRIs can change the strategic orientation to participate in conferences or networks for identifying suitable partners. Followed strategies by PRIs can be evaluated and adapted over time. The model elements that have been translated into the EIS simulation model are in bold

innovations are relatively limited. When full investments are made, PRIs perform strong, in particular regarding the amount of collaborations and innovations produced. Improving in only exploration processes does not improve the overall innovation performance of PRIs, as it does not positively influence a PRI's more downstream activities.

Participate in conferences: When the cost of conferences is low, it is almost always beneficial to go to conferences due to an increased likelihood of finding potential partners for collaboration. When the cost increases, participating in conferences becomes less successful compared with non-participation. However, when the frequency of participation in conferences is reduced, the fitness and innovation performance of PRIs can be improved. This seems to relate to the spreading of risks and existing trade-offs between costs and finding partners. In practice, this could be related to institutes picking selected conferences in which to participate, even when there are more conferences that could be potentially relevant.

Participate in networks: Institutes that participate in networks are often successful, especially when the costs for participating in networks are low. As expected, an increased cost of network participation diminishes the beneficial effect of networks. However, when networks are only partially utilised by PRIs, the overall fitness of PRIs can be high, even with relatively high networking costs. This is a similar pattern as observed for conferences.

A graphical illustration of the simulated scenarios is provided in Fig. 44.3.

44.4 Discussion and Outlook

To model enabling innovation within PRIs, the combination of an agent-based with a system-dynamics dimension allows the formalised addressing of relevant (inter)organisational processes and capabilities for assessing PRI collaboration strategies. Although the calibration of the initiated EIS model remains a daunting task to be accomplished yet – particularly due to the difficult accessibility of sensitive organisational data – the simulated scenarios have provided some interesting insights. For example, the model shows that investing in specific organisational capabilities for improving collaboration with target groups leads to an overall better innovation performance of PRIs. This is particularly evident when PRIs take an integrative approach, that is, invest in all explorative, exploitative and evaluative processes to initiate and manage collaboration. Also, both conferences and networks can be very beneficial platforms for PRIs for establishing fruitful collaborations. However, PRIs should consider closely the investments that need to be made when using these platforms and the frequency these platforms should be exploited.

In order to make the model more practice-based and suited for customised simulation, several steps can be taken for future research. These include the collection of appropriate organisational data, empirical-based calibration and validation of the model, the matching of the empirical basis to the model architecture through a legitimate indicator system, further expansion of potential direct and indirect effects by the modelled agents and the inclusion of a richer environment to allow improved contextualised simulation and scenario building. Such steps could bring the model closer to becoming a valuable tool for PRIs that aim to be enablers of innovation.

Fig. 44.3 Results *Collaboration* scenarios. Simulation preliminary results. Top row = size populations or focus population following strategy within scenario (red) and control population (blue); Bottom row = amount of collaborations (red = focus population, green = control population) and amount of innovations (blue = focus population, purple = control population). To assess the

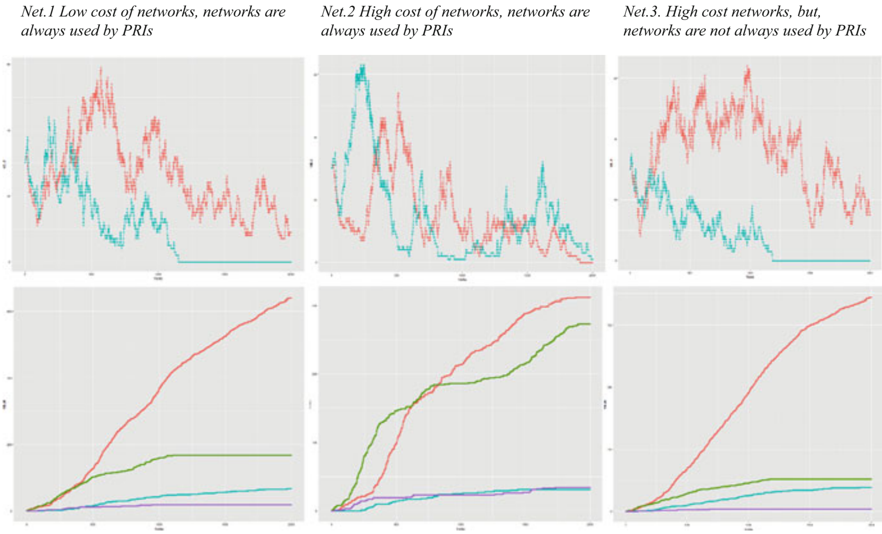


Fig. 44.3 (continued) viability of specific strategy configuration under certain environmental conditions, two populations (with equal number of agents) are created from a certain agent type (i.e. having a certain strategy configuration). For the conducted simulations, one of these population follows a specific strategy (focus population) – that is, has a configuration of certain organisational variables that drive the specific strategy during the total run of the simulation. The other population does not follow this specific strategy and therefore functions as a control group (control population). The success of a strategy is determined based on the amount of generated resources, innovations and spin-offs by the two different populations. Scenarios were repeated a number of times to check the consistency of model output. As the architecture of the model remains relatively simple, it is still susceptible to some of the random and uncontrolled variation in agent behaviour. This leads sometimes to different outcomes of repetitions having the same initialization and modelling conditions. Nevertheless, for all scenarios, a dominant pattern could be distinguished. The graphs depicted here are selected run repetitions of the defined scenarios that represent these scenario-specific patterns

Acknowledgement The EIS-project was funded by the German Federal Ministry of Education and Research (BMBF – funding ID FKZ: 01IO1509).

References

1. Gibbons, M., Johnston, R.: The roles of science in technological innovation. Res. Policy. **3**, 220–242 (1974)
2. Schlaile, M.P., Ehrenberger, M.: Complexity in Entrepreneurship, Innovation and Technology Research. Springer, Cham (2016)
3. von Hippel, E.: The Sources of Innovation. MIT Press, Cambridge, MA (1988)
4. Kline, S.J., Rosenberg, N.: An overview of innovation. Eur. J. Innov. Manag. **38**, 275–305 (1986)
5. Nelson, R.R.: Capitalism as an engine of progress. Res. Policy. **19**, 193–214 (1990)

6. Cohen, W.M., Nelson, R.R., Walsh, J.P.: Links and impacts: the influence of public research on Industrial R & D. Manag. Sci. **48**, 1–23 (2002)
7. Nadler, D.A., Tushman, M.L.: A model for diagnosing organizational behavior. Organ. Dyn. **9**, 35–51 (1980)
8. Zajac, E.J., Kraatz, M.S., Bresser, R.K.F.: Modeling the dynamics of strategic fit: a normative approach to strategic change. Strateg. Manag. J. **21**, 429–453 (2000)
9. Gerhards, U., Köller, C., Faix, A., Volkmann, M.: «Enabling Innovation". Wissens- und Technologietransfer – Pflicht und Kür für Forschungseinrichtungen. Wissenschaftsmanagement(2012)
10. Duffy, J.: Chapter 19 Agent-based models and human subject experiments. In: Tesfatsion, L., Judd, K.L.(eds.) Handbook of Computational Economics, vol. 2, pp. 949–1011. Elsevier (2006)
11. Kohler, T.A., Gumerman, G.J., Reynolds, R.G.: Simulating ancient societies. Sci. Am. **293**, 76–84 (2005)
12. Gilbert, N., Troitzsch, K.G.: Simulation for the Social Scientist. McGraw- Hill Education (UK), London (2005)
13. Goldstein, J.: Emergence as a construct: history and issues. Emergence. **1**, 49–72 (1999)
14. Pyka, A., Grebel, T.: Agent-based modelling – a methodology for the analysis of qualitative development processes. In: Agent-Based Computational Modelling: Applications in Demography, Social, Economic and Environmental Sciences (Contributions to Economics), pp. 17–35 (2006)
15. Edquist, C. (ed.): Systems of Innovation: Technologies, Institutions and Organizations. Routledge, London (1997)
16. Ghorbani, A., Dijkema, G., Schrauwen, N.: Structuring qualitative data for agent-based modelling. JASSS. **18**, 1–6 (2015). http://jasss.soc.surrey.ac.uk/JASSS.html

Chapter 45
Using Social Simulations in Interdisciplinary Primary Education: An Expert Appraisal

Pauline Vos, Markos Dallas, Amrit B. Poudel, and F. LeRon Shults

Abstract Many people recognize that teaching basic skills in primary schools (reading, writing, and arithmetic) is no longer sufficient for pupils in the digital age. Therefore, governments now increasingly ask schools to add other skills (oral, digital) and to create connections between subjects (e.g., use mathematics in history lessons). In this study, we explored how social simulations can be used in primary education to meet these new goals. We conducted an expert appraisal (a qualitative Delphi method) with four experts specializing in innovating primary education. We selected three simulations that were freely available on the web, relevant for pupils' lives and had a limited number of parameters. They dealt with segregation, gossip spread and population dynamics. We asked the experts to critically discuss these. Afterward, we analyzed the videotaped discussions in terms of affordances and constraints. The results showed that the affordances of social simulations include their broad appeal to students and their capacity to help users explore relevant concerns through an integrative approach (e.g., interpreting graphs, reasoning with parameters, predicting). Also, the experts warned that social simulations can touch on ethical issues that might be stressful for some pupils. If well-orchestrated, the use of social simulations has great potential to fulfill the new primary school goals.

Keywords Expert appraisal · Interdisciplinary primary education · Social simulations in schools

45.1 Introduction

In primary schools all over the world, pupils learn to read, write, and calculate. These three basic skills were considered important when compulsory education was established. More than a century later, many people question whether these are still

P. Vos (✉) · M. Dallas · A. B. Poudel · F. LeRon Shults
University of Agder, Kristiansand S, Norway
e-mail: pauline.vos@uia.no; markos.dallas@uia.no; amrit.poudel@uia.no; leron.shults@uia.no

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_45

sufficient and whether other skills (reasoning, presenting, visualizing data, computing, etc.) should be added. Moreover, some have observed that compartmentalizing the teaching of these basic skills does not foster the emergence of higher-order reasoning required for complex problems [1]. Such compartmentalization can also create dissonance because people seem to "forget" their skills when confronted with out-of-school problems [2].

In many countries, educational ministries are trying to remedy this situation and improve learning outcomes. The Norwegian Ministry of Education and Research [3] has framed five basic skills (oral skills, reading, writing, digital skills, and numeracy) and requires these to be intertwined in teaching. For example, visual-numeric representations (e.g., bar charts) should be part of social subjects (history, geography, and religious education). Also, learning experiences should be interactive (hands-on), related to out-of-school life, and ask pupils for higher-order reasoning and solving complex problems. Against this background, we were inspired by the decades of work by Uri Wilensky and colleagues [4, 5], who demonstrated that pupil's interaction with digital simulations of complex systems offers valuable learning experiences. However, their focus is on (1) "learning to code" (e.g., making robots move) and (2) connecting simulations to learning concepts of the natural sciences. Alternatively, we focus on (1) students' meta-knowledge (knowing about) of simulations, which entails understanding aims, limitations, and when/where/for what purposes they can be used (or not), and (2) simulations of social phenomena.

Social simulations can connect themes from pupil's lives to representations of change (e.g., simulating how a group of kids share news and how the spread increases and fades). We hypothesized that through social simulations, pupils' learning experiences could meet the government's educational goals. Therefore, we started a study on the use of social simulations with 10–13-year-old pupils. Applying a sociocultural perspective, we framed pupils as actors within (1) school and (2) out-of-school activities. In schools, pupils' activities are prompted by designed tasks, and simulations can be tools, which have affordances (what it offers and provides, subjectively perceived by a user) and constraints (the conditions which guide and limit the activities) [6]. Our research question was: what are affordances and constraints in using social simulations in primary education to meet the government's goals?

45.2 Materials and Participants

We organized an expert appraisal (a qualitative Delphi method) in a session of 3 hours. The program entailed: (1) introduction and goals of the project, (2) viewing and critically discussing three selected digital simulations, and (3) a general discussion. We selected three agent-based simulations on the criteria: freely available on the web, connecting to pupil's lives, and having a limited number of parameters (sliders).

1. A simulation of Schelling's segregation model with four sliders (similar, red/blue, empty, size), which can connect to pupils' experience of peer group clusters and the inclusion or exclusion of people; see Fig. 45.1 (top left). Available from http://nifty.stanford.edu/2014/mccown-schelling-model-segregation/.
2. A gossip simulation with three sliders (speed, message loss, number of initial messages), which connects to how information spreads among a group; see Fig. 45.1 (top right), http://brucespang.com/bimodal/simulation/.
3. A population dynamics simulation with three graphical representations (a social network, a population tree, and a time graph), one slider (fertility); see Fig. 45.1 (bottom). The simulation can connect to the pupils' experience of population growth and family life. Available from https://runthemodel.com/models/1633/.

Four experts (two women, two men) participated; they will be indicated as P1–P4. They were selected because of their educational experience with (1) the age group, (2) educational use of digital tools, (3) inquiry-based learning, and (4) innovating education. The discussions were open and only slightly guided to keep the focus on the goals of the project. The experts sat together with one laptop, on which their activities were screen-captured. Additionally, we filmed facial expressions and gestures. We analyzed transcripts on affordances and constraints of including the simulations into primary education. The affordances and constraints pertained to activities within school (using, reasoning, and understanding) and to connecting to pupil's out-of-school life.

45.3 Results

Space only allows us to summarize results. First, we found that the experts needed more time than anticipated (approx. 30′) to familiarize themselves with the simulations. This revealed a constraint for first-time users. As for the number of variables (100 or 900 agents in a simulation), the participants did not appear to be thrown off by higher numbers. However, a simulation with only one slider was considered too simple (one can easily see its effect), two sliders were more challenging, while a simulation with four sliders was considered complex and required more guidance. Another constraint was the abstractness of parameters: for example, fertility would better connect to pupil's imagination if expressed as "average number of children per woman" rather than as "number of births per 1000 inhabitants." As for the title of the simulations, they rather had the gossip simulation have the name "information spread simulation" as the term gossip has a negative connotation (slandering). Regarding the interface, the experts preferred one field showing the agents moving or changing color; multiple graphical representations within one window was considered a constraint.

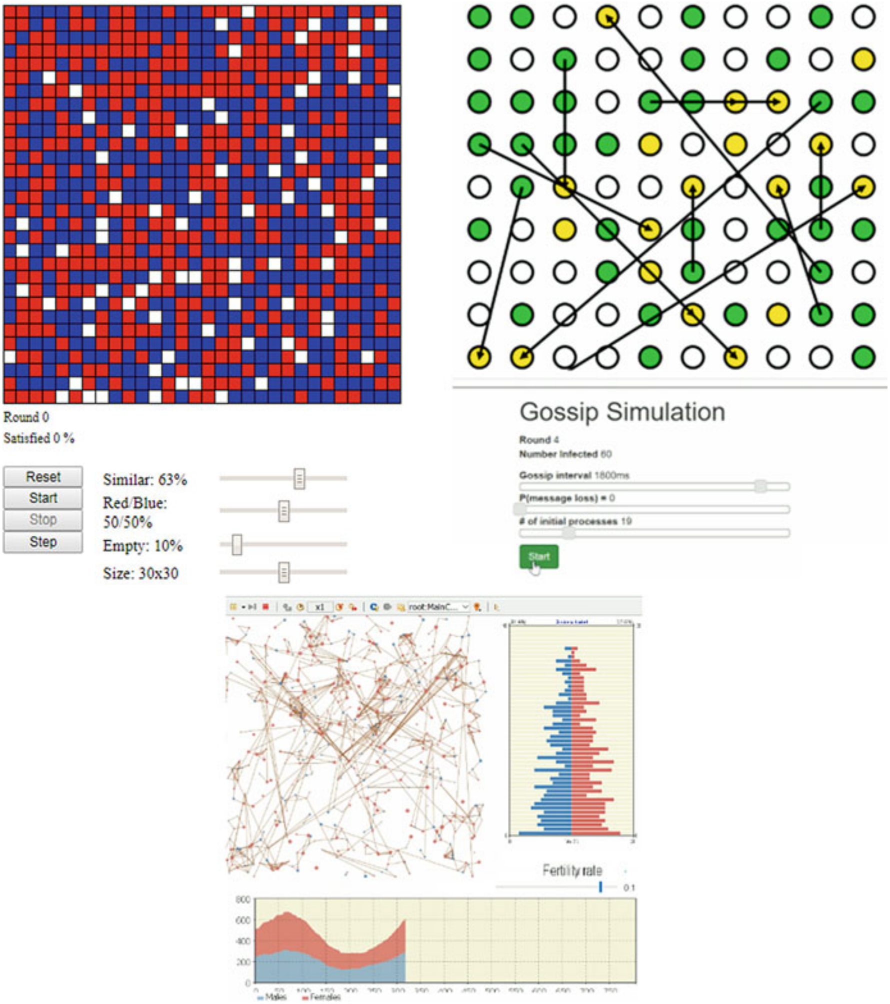


Fig. 45.1 The interface of the segregation simulation (top left), the gossip simulation (top right), and the population dynamics simulation (bottom)

The simulations afforded a connection between pupil's lives and the social subjects in schools. The experts discussed whether a clear connection to mathematics could be made, since mathematics is *black boxed* [7] in simulations (in part because sliders are more user-friendly handles for parameters). Thus, traditional school mathematics is not visible at a surface level in the simulations. However, P3 observed that any simulation has a time scale, which allows for tasks that ask for quantitative reasoning and making predictions (e.g., "how long does it take to . . . ?"). Also, the experts noted that the simulations offer opportunities for interpreting visual representations of change; for example, pupils' reasoning about a "decreasing increase" is a precursor to intuitive understanding of the second derivative. Thus, the simulations afforded interdisciplinary thinking (combining social sciences and quantitative reasoning).

The experts also considered social simulations to be appealing to a wide group of pupils (girls, pupils from disadvantaged backgrounds, etc.), reaching beyond the pupils who have an interest in science and technology. They indicated a preference for the "lighter" themes of information spread and fertility over themes such as segregation, considering the age of the pupils. The latter theme triggered discussion:

P3: But, what in a multicultural classroom some of them will not be motivated at all (..) It will be confronting.
P1: Yes, also if it is with neighbors rich and poor.
P3: Maybe, ... introduce first something which is safe like ... hm ... are you supporter of this soccer team or that soccer team or ... (..) If you look how many do a certain sport ... if 60% do the same ...

The experts worried that themes touching on social or racial discrimination could wound pupils. They argued that teachers should be aware of delicate issues, use care when handling ethical themes in classrooms, or begin with less controversial themes, for example, by replacing racial segregation by rivalry between sports fans.

45.4 Conclusions, Discussion, and Recommendations

In this study, we explored some of the affordances and constraints related to the use of social simulations in primary education. Through an expert appraisal, we found that social simulations afford the creation of connections between pupils' out-of-school experiences and school tasks and between different school subjects (social subjects, mathematics). An overarching value of adding social simulation to the curriculum at primary schools is that they are appealing to many pupils, including underrepresented groups in science and technology. However, the experts also identified some potential constraints: the usability of the tools (e.g., ease of understanding the interface), the need for didactical task design (e.g., connecting the subject of mathematics to simulations), and the importance of carefully handling ethical themes (e.g., discussing segregation could be risky for some pupils).

We should use caution in interpreting these results because our study was small-scale and explorative. Clearly, more research is needed especially into the way in which simulation could be used to facilitate higher-order thinking, how to scaffold activities, and what transfer takes place from understanding simulations to understanding real-world scenarios. We argue that social simulations can support the government's educational goals, but their implementation will require careful orchestration. Since primary education is still currently compartmentalized, teachers will need assistance and professional development as they make the transition and explore new integrative pedagogical approaches.

References

1. Mitra, S., Dangwal, R., Chatterjee, S., Jha, S., Bisht, R.S., Kapur, P.: Acquisition of computing literacy on shared public computers: children and the "hole in the wall". Australas. J. Educ. Technol. **21**, 407–426 (2005)
2. National Research Council: How People Learn: Brain, Mind, Experience, and School. National Academies Press, Washington, D.C. (2000)
3. Norwegian Ministry of Education and Research: Framework for Basic Skills. UDIR, Oslo (2012)
4. Wilkerson-Jerde, M.H., Wilensky, U.J.: Patterns, probabilities, and people: making sense of quantitative change in complex systems. J. Learn. Sci. **24**, 204–251 (2015)
5. Brady, C., Orton, K., Weintrop, D., Anton, G., Rodriguez, S., Wilensky, U.: All roads lead to computing: making, participatory simulations, and social computing as pathways to computer science. IEEE Trans. Educ. **60**, 59–66 (2017)
6. Kennewell, S.: Using affordances and constraints to evaluate the use of information and communications technology in teaching and learning. J. Inform. Techol. Teach. Educ. **10**, 101–116 (2001)
7. Williams, J., Wake, G.: Black boxes in workplace mathematics. Educ. Stud. Math. **64**, 317–343 (2007)

Chapter 46
Switching Costs in Turbulent Task Environments

Friederike Wall

Abstract Altering the status quo in favor of a new solution may induce particular extra costs. In organizations such switching costs are assumed to prevent decision-makers from implementing new solutions and, hence, to cause inertia. Particularly when organizations operate in a turbulent environment which requires determined adaptive behavior, switching costs are suspected to unfold detrimental effects on the performance of an organization. This paper studies the effects of switching costs for different levels of turbulence in the task environment and controls for the complexity of interactions within an organization. For this, an agent-based simulation based on NK fitness landscapes is employed. The results indicate that the effects of switching costs subtly vary with the complexity and the turbulence of the task environment. Moreover, contrary to the prevailing view, the study leads to the conjecture that switching costs may increase efficiency of adaptive processes particularly in turbulent and complex task environments.

Keywords Agent-based simulation · Complexity · Cost of effort · NK fitness landscapes · Turbulence

46.1 Introduction

In various domains related to organizational thinking, it is well recognized that changing the status quo in favor of an alternative may induce one-time costs (e.g., [1–3]). These costs may comprise, for example, of cognitive effort for learning, opportunity costs due to the scarce attention of management during reorganizations,

F. Wall (✉)
Alpen-Adria-Universitaet Klagenfurt, Klagenfurt, Austria
e-mail: friederike.wall@aau.at

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_46

costs for the changeover of technical equipment, costs for dealing with resistance of certain stakeholders, or (emotional) costs for loosing old and building new relations (e.g., [4–7]). These costs often are captured by the term "switching costs."[1]

While it is well recognized that switching costs occurring on the site of market partners may have their positive effects for a firm (e.g., in customer retention strategies, [10]), they are suspected to have detrimental effects when occurring within the organization. In particular, a reasonable conjecture is that switching costs reduce the propensity that new and, potentially, superior solutions are implemented and, thus, may cause that organizations stay with the status quo instead of changing in favor of the new [1]. Hence, switching costs often are regarded as one of the inertial pressures on organizations besides, for example, the complexity of the task to be performed by an organization. From an organizational ecology perspective, it has been argued that inertia might be particularly devastating in turbulent environments (e.g., [11, 12]). However, not only is the subtle interrelation between inertia and adaptability a field of ongoing research (e.g., [12–14]); rather, has the particular role of switching costs in the field of organizational inertia and adaptation to environmental change, so far, received little attention.

Against this background, the paper seeks to contribute to the following research question: *Which effects on organizational performance result from switching costs in turbulent task environments?* However, not only switching costs but also the complexity of the overall task of an organization may be a predominant inertial force in organizations (e.g., [1]); therefore, the study also takes complexity of the task environment as contextual factor into account.

For investigating the research question, the paper makes use of an agent-based simulation. A simulation-based research method appears appropriate to capture a longitudinal perspective which is required for studying environmental turbulence and to systematically control for different levels of task complexity and switching costs. An agent-based simulation allows to consider the collaboration of various interacting parties (e.g., units) within an organization [15]. In the model, the task environment of the organizations is represented according to the framework of NK fitness landscapes [16, 17] which was originally introduced in the domain of evolutionary biology and, since then, broadly employed in managerial science [18]. A key feature of the NK framework is that it allows to easily control for the complexity of the task to be performed [19].

In the remainder of the paper, after the simulation model is outlined, Sect. 46.3 gives an overview over the simulation experiments, and in Sect. 46.4 results are introduced and discussed.

[1] In a stricter sense the term "switching costs" is used in the domain of strategic management in the context of switching from one provider to another; see esp. [8, 9].

46.2 Simulation Model

46.2.1 Task Environment: Complexity and Turbulence

In the simulations, artificial organizations are observed while searching for superior solutions for a task to be accomplished. The task's characteristics are captured according to the framework of NK fitness landscapes with its particular capability to control for the complexity of the task [16, 17]. Moreover, the organizations may operate in a stable or in a turbulent environment. In the latter case, the task environment undergoes more or less radical exogenous shocks.

In particular, following the NK framework, the organizations face an N-dimensional binary decision problem $\mathbf{d_t} = (d_{1t}, \ldots, d_{Nt})$ with $d_{it} \in \{0, 1\}$, $i = 1, \ldots N$. These N binary choices may reflect, for example, whether only in-house production or also outsourcing is employed, whether a product is included in the product portfolio or excluded, or whether the distribution of products is done via shops or not, via the Internet, and so on. In every time step t, each of the two states of each single choice d_{it} of the N-dimensional binary problem $\mathbf{d_t}$ contributes with C_{it} to the overall performance V_t of the organization. The overall performance V_t achieved in period t is defined as the normalized sum of contributions C_{it}:

$$V_t = V(\mathbf{d_t}) = \frac{1}{N} \sum_{i=1}^{N} C_{it} \qquad (46.1)$$

The key feature of the NK framework is that it allows to control for the overall task's complexity captured by parameter K. In particular, the contributions C_{it} can be subject to interactions among the N single choices. Parameter K (with $0 \leq K \leq N - 1$) gives the number of choices d_{jt}, $j \neq i$ which also affect the performance contribution C_{it} of choice d_{it}. (For the sake of simplicity, as in this paper, K often is the same for each single choice d_{it}.) Hence, contribution C_{it} is not only affected by the choice d_{it}, but may also depend on K other choices:

$$C_{it} = f_i(d_{it}; d_{i_1 t}, \ldots d_{i_K t}) \qquad (46.2)$$

with $\{i_1, \ldots, i_K\} \subset \{1, \ldots, i-1, i+1, \ldots, N\}$. Thus, task complexity is at the lowest level with $K = 0$, when choices do not affect each other; if each single choice i affects the performance contribution of every other choice $j \neq i$, complexity is at maximum with $K = N - 1$. The C_{it} are randomly drawn from a uniform distribution with $0 \leq C_{it} \leq 1$. With this, from a more "technical" point of view, NK landscapes are stochastically generated pseudo-boolean functions with N bits, i.e., $F : \{0, 1\}^N \rightarrow \mathbb{R}^+$, [19]. The number of different C_{it} to be generated is given by $N \cdot 2^{K+1}$, and, together, the C_{it} form the "performance landscape" (corresponding to the original term "fitness landscape" [16, 17]).

The model employed in this study distinguishes between stable and turbulent task environments. In case of stability, the contributions C_{it} remain unchanged over the entire observation period T after the performance landscape once has been computed in the beginning of the observation, i.e., the simulation run. In contrast, as in [13], turbulence is captured by exogenous "correlated shocks" which occur from time to time and change the performance landscape. In particular, for organizations operating in a turbulent task environment, periodically, after each T^*-th period, a new performance landscape is computed while keeping the pattern of interactions among the single choices i and, thus, the same level of K. The new landscape results as a "mixture" of the old and the newly generated performance contributions (also drawn from a uniform distribution following the functional relation denoted in Eq. 46.2). Parameter θ (with $0 \leq \theta \leq 1$) captures the ratio to which the "old" environment is represented in the new landscape:

$$C_{it} = \theta \cdot C_{i,(n(t)-1)\cdot T^*} + (1-\theta) \cdot u_{n(t)\cdot T^*} \tag{46.3}$$

for $n(t) \cdot T^* + 1 < t \leq (n(t-1)+1) \cdot T^*$ where $n(t) = \lfloor t/T^* \rfloor$ gives the number of shocks that, so far, have occurred in the observation period. While keeping the pattern of interactions among the choices d_i, u is a new draw from a uniform distribution over the unit interval performed after period $n(t) \cdot T^*$.

Hence, according to Eq. 46.3, the higher θ the higher is the correlation between the old and the new performance landscape; or putting it the other way round, the lower θ the higher the turbulence: with $\theta = 0$ the new landscape is not affected by the previous one, and if $\theta = 1$, there is, in fact, no shock.

46.2.2 Organizational Structure, Units, and Switching Costs

The N-dimensional problem is decomposed into M disjoint and equal-sized subproblems, indexed by $r = 1, \ldots, M$ and with size $N^r = N/M$. Each of the M subproblems is exclusively assigned to one unit r within the organization, and each unit searches for superior solutions for its N^r-dimensional sub-task. There is no intervention by any central decision-making authority, and the headquarter is confined to observe the performance achieved in the end of each period.

With this, the units serve as the agents whose decisions drive an organization's adaptive walk on the performance landscape. In particular, in every time step t, each unit has to decide which configuration $\mathbf{d^r_t}$ of the subproblem assigned to that particular unit is to be implemented. However, in order to capture limited information-processing capabilities of decision-making agents as is familiar in agent-based modeling [18], we restrict the number of alternatives that each unit can choose from. In particular, each unit can choose one out of three options: first, a unit may opt for keeping the status quo $\mathbf{d^{r*}_{t-1}}$, i.e., the option chosen in the previous period; second, by random, each unit discovers an alternative $\mathbf{d^{r,a1}_t}$ that differs in one of the binary choices of the status quo-configuration, i.e., where the Hamming distance $h(\mathbf{d^{r*}_{t-1}}, \mathbf{d^{r,a1}_{t}}) = \sum_{i=1}^{N^r} \left|\mathbf{d^{r*}_{t-1}} - \mathbf{d^{r,a1}_{t}}\right|$ equals 1; third, each unit identifies an option $\mathbf{d^{r,a2}_{t}}$ where two single choices are altered compared to the status quo.

In the model, the number of bits flipped is regarded as the *effort* to be taken by unit r in order to implement the chosen option and, with this, provides the basis for capturing the switching costs. Hence, the effort – given by the Hamming distance – is

$$h(\mathbf{d^{r*}_{t-1}}, \mathbf{d^{r}_{t}}) = \begin{cases} 0 \text{ if } \mathbf{d^{r}_{t}} = \mathbf{d^{r*}_{t-1}} \\ 1 \text{ if } \mathbf{d^{r}_{t}} = \mathbf{d^{r,a1}_{t}} \\ 2 \text{ if } \mathbf{d^{r}_{t}} = \mathbf{d^{r,a2}_{t}} \end{cases} \tag{46.4}$$

For modeling the costs of effort, as usual (e.g., [20, 21]), it is assumed that higher levels of effort are increasingly costly. Hence, for unit r's costs of effort S^r, we have $S^r(h)' > 0$ and $S^r(h)'' > 0$. In particular, the switching costs $S^r(\mathbf{d^{r}_{t}})$ of unit r are modeled to be quadratically increasing with h, i.e.,

$$S^r_t(\mathbf{d^{r}_{t}}) = s^r \cdot (h(\mathbf{d^{r*}_{t-1}}, \mathbf{d^{r}_{t}}))^2 \tag{46.5}$$

where s^r is a cost-coefficient. For the sake of simplicity, in the simulations the cost-coefficient is the same for all units r and does not change over time.

46.2.3 Units' Decision-Making

Each unit r pursues its "own" objective – related to its respective subproblem. In particular, each unit seeks to identify that option out of $\mathbf{d^{r*}_{t-1}}$, $\mathbf{d^{r,a1}_{t}}$ and $\mathbf{d^{r,a2}_{t}}$ which promises the highest net performance A^r, i.e., the highest difference between partial performance P^r_t – resulting from the contributions C_{it} of those particular single choices assigned to unit r – and the related switching costs

$$A^r_t(\mathbf{d^{r}_{t}}) = P^r_t(\mathbf{d^{r}_{t}}) - S^r_t(\mathbf{d^{r}_{t}}) \tag{46.6}$$

with the partial performance P^r_t given by

$$P^r_t(\mathbf{d^{r}_{t}}) = \frac{1}{N} \sum_{i=1+(r-1)N^r}^{rN^r} C_{it} \tag{46.7}$$

with $N^r = N/M \forall r$. Hence, unit r is focused on the net performance obtained from that partial problem $\mathbf{d^{r}_{t}}$ which is in its own primary control and seeks to decide in favor of that option which promises the highest net performance A^r_t.

However, when ex ante evaluating the performance P_t^r of options and, thus, forming preferences, each unit r suffers from two kinds of imperfect information: First, unit r does not know the intentions of the other units $q \neq r$; rather, unit r assumes that they will stay with their previous choices which, obviously, is of particular relevance, when interactions among the units' subproblems occur. Second, unit r cannot precisely foresee its options' effects on partial performance P_t^r. In particular, the model captures distortions from the true consequences of options, and these are depicted as relative errors z^r ([22]; for other functions see [23]) where the z^r follow a Gaussian distribution $N(0;\sigma)$ with expected values 0 and standard deviations σ^r. For simplicity's sake, σ^r is the same for each unit r; errors are assumed to be independent from each other. Hence, rather than maximizing Eq. 46.6, unit r, in fact, maximizes the *perceived* net performance:

$$\widetilde{A_t^r}(\mathbf{d_t^r}) = \widetilde{P_t^r}(\mathbf{d_t^r}) - S_t^r(\mathbf{d_t^r}) \tag{46.8}$$

where

$$\widetilde{P_t^r}(\mathbf{d_t^r}) = P_t^r(\mathbf{d_t^r}) + z^r(\mathbf{d_t^r}) \tag{46.9}$$

46.3 Simulation Experiments

The simulation experiments are intended to study the effects of switching costs in differently complex and turbulent task environments. Parameter settings are displayed in Table 46.1. In the simulations, artificial organizations are "thrown" randomly into a performance landscape showing one of four patterns of interactions (a, b1 to b3 in Table 46.1) and observed over $T = 400$ periods while searching

Table 46.1 Parameter settings

Parameter	Values/types
Observation period	$T = 400$
Number of choices	$N = 12$
Number of agents	$M = 4$ with $\mathbf{d^1} = (d_1, d_2, d_3)$, $\mathbf{d^2} = (d_4, d_5, d_6)$, $\mathbf{d^3} = (d_7, d_8, d_9)$, $\mathbf{d^4} = (d_{10}, d_{11}, d_{12})$
Interaction structures	a. decomposable: $K = 2$, $K^{ex} = 0$; b. non-decomposable: b1: $K = 3$, $K^{ex} = 1$; b2: $K = 5$, $K^{ex} = 3$; b3: $K = 7$, $K^{ex} = 5$
Precision of ex ante evaluation	$\sigma^r = 0.1 \;\; \forall r$
Cost-coefficient	$s^r \in \{0, 0.02\} \, \forall r$
Interval between shocks	$T^* = 100$
Correlation of landscapes after shock	$\theta \in \{0, 0.25, 0.5, 0.75, 1\}$

for superior levels of performance for an $N = 12$-dimensional problem. For each simulation run, in time step $t = 0$, an initial performance landscape is generated, and the individually distorted views of the $M = 4$ units are computed. After every $T^* = 100$-th period, an external shock occurs which affects the performance landscape (and the respective distorted units' views) according to the level θ of correlation between the old and the new performance landscape. For figuring out the effects of switching costs, the experiments pursue the idea of a factorial design [24] and, based on some pretests, distinguish between scenarios where no switching costs ($s^r = 0$) occur and the respective scenarios with non-zero switching costs ($s^r = 0.02$).[2]

The experiments distinguish for different levels of task complexity. In particular, two parameters are employed for capturing the pattern of interactions among the units of the artificial organizations: While parameter K reflects the overall level of complexity (see Sect. 46.2.1), K^{ex} denotes the level of interactions *across subproblems* and, with that, also reflects the interactions among the units. The simulation results are obtained for different interaction structures characterized by K and K^{ex} for $N = 12$ (Table 46.1):

(a) In a perfectly decomposable structure an organization's overall decision problem, $\mathbf{d_t}$ is fragmented into $M = 4$ disjoint parts with maximal intense intra-subproblem interactions, but no cross-subproblem interactions (i.e., $K = 3, K^{ex} = 0$) as sketched in Fig. 46.1a. This reflects a block-diagonal

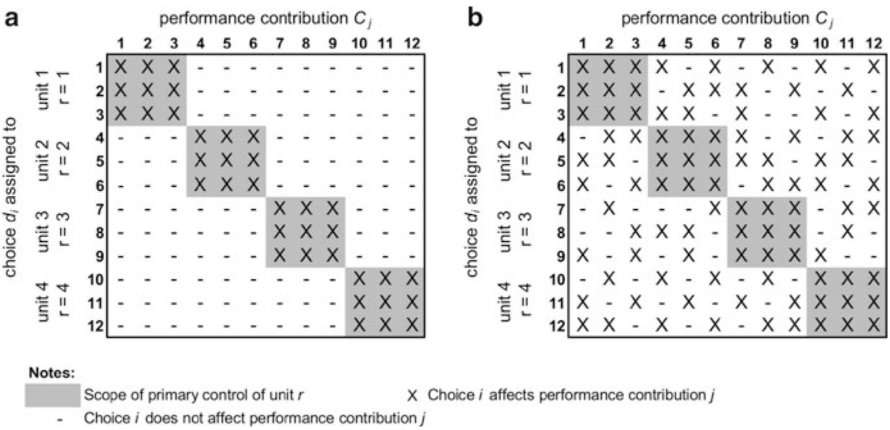


Fig. 46.1 Examples of interaction structures and assignment of choices to units. (**a**) decomposable structure ($K=2$, $K^{ex}=0$): no cross-unit interactions (**b**) non-decomposable structure ($K=7$, $K^{ex}=5$): examplary medium level of cross-unit interactions

[2]For a more detailed analysis of different levels of switching costs in stable environments, see [25]; given the ranges of the C_{it} between 0 and 1 and the units' objective functions in Eqs. 46.6 and 46.7, $s^r = 0.02$ reflects a rather high level of switching costs.

structure [26] where the decision problem can be decomposed into disjoint partial problems, and each of these is delegated to one distinct unit. This corresponds to a "self-contained" organization structure [27] and comes close to a pooled interdependence [28]. These structures are particularly feasible for organizations pursuing strategies of geographical diversification or which create their units according to products in terms of product divisions.

(b) In non-decomposable structures with $K^{ex} > 0$, interactions across the subproblems occur. Hence, performance contributions C_i – though in primary control of unit r – may be affected by some choices made by other units $q \neq r$. For example, Fig. 46.1b could represent an organization with a functional specialization with its typically high level of interrelations between units. This situation comes close to what, according to Thompson's classification, is called reciprocal interdependencies [28].

46.4 Results and Discussion

The results of the simulation experiments are introduced in two steps: Sect. 46.4.1 presents results obtained for baseline scenarios in order to gain a basic understanding of effects of switching costs for selected levels of task turbulence and complexity; and Sect. 46.4.2 studies the sensitivity of results for a broader range of environmental turbulence and complexity.

46.4.1 Analysis of Baseline Scenarios

According to the idea of factorial design [24], results obtained for two interaction structures (i.e., the decomposable and the non-decomposable according to Fig. 46.1a, b, respectively) with zero and high switching costs ($s^r = 0$ vs. $s^r = 0.02$) are reported – each for a stable and a moderately turbulent task environment ($\theta = 1$ vs. $\theta = 0.75$). Hence, eight baseline scenarios arise for which Table 46.2 displays the respective results: each row represents the results obtained for 500 simulation runs with the parameter settings indicated in the leftmost column (a). In particular, the final performance[3] (col. b) and the performance achieved on average over the 400 periods, $\overline{V}_{(0;400)}$ (col. c), inform about the effectiveness of the adaptive walks performed by the organizations. The ratio of periods in which a new configuration $\mathbf{d_t}$ is implemented (d) and the relative frequency of periods where the status quo is altered in favor of false-positive options from the organization's perspective (e) as

[3]The final performance $V_{t=400}$ and the average performance $\overline{V}_{(0;400)}$ are given in relation to the global maximum of the respective performance landscapes: otherwise the results could not be compared across different performance landscapes.

Table 46.2 Condensed results for the baseline scenarios. Each row represents results of 500 simulation runs, each on distinct performance landscapes. For parameter settings, see Table 46.1

Scenario of costs and turbulence (a)	Final perf. $V_{t=400}$ ±CI* (b)	Average perf. $\overline{V}_{(0;400)}$ ±CI* (c)	Freq.New Conf. $\mathbf{d_t}$ (d)	Freq. FalsePos. (e)	Av.Eff. $\overline{h}_{(0;400)}$ (f)
Decomposable structure with $K = 2, K^{ex} = 0$					
$s^r = 0, \theta = 1$	0.9887 ±0.0024	0.9876 ±0.0021	21.1%	10.2%	4.6%
$s^r = 0.02, \theta = 1$	0.9328 ±0.0074	0.9319 ±0.0073	1.0%	0.2%	0.2%
$s^r = 0, \theta = 0.75$	0.9841 ±0.0025	0.9850 ±0.0012	32.9%	16.2%	7.8%
$s^r = 0.02, \theta = 0.75$	0.9558 ±0.0051	0.9470 ±0.0050	2.1%	0.2%	0.2%
Non-decomposable structure with $K = 7, K^{ex} = 5$					
$s^r = 0, \theta = 1$	0.7416 ±0.0224	0.7119 ±0.0111	81.3%	41.1%	38.2%
$s^r = 0.02, \theta = 1$	0.8454 ±0.0110	0.8396 ±0.0106	3.8%	1.7%	0.8%
$s^r = 0, \theta = 0.75$	0.7896 ±0.0120	0.7394 ±0.0043	92.3%	46.5%	43.8%
$s^r = 0.02, \theta = 0.75$	0.8862 ±0.0084	0.8607 ±0.0065	7.2%	3.5%	1.2%

*Confidence intervals at a confidence level of 0.99

well as the relative effort[4] that was made in sum by all $M = 4$ units on average per period (f) characterize the adaptive search processes more into detail. Figure 46.2 plots the respective adaptive walks of the eight baseline scenarios, organized by the interaction structure.

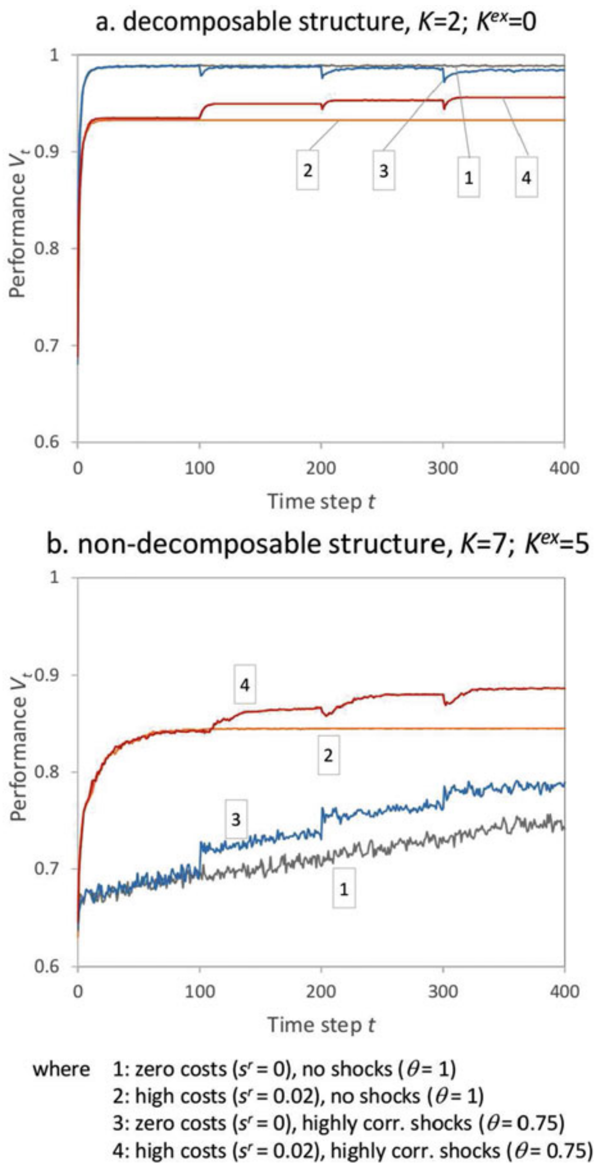
The results broadly support the conjecture that with switching costs alterations from the status quo become less frequent. As can be seen in Table 46.2, all else equal, the frequency of alterations in $\mathbf{d_t}$ (col. d) in all of the four scenarios with non-zero switching costs is remarkably lower than in the respective scenarios with costless alterations; correspondingly, the average relative effort per period (col. f.) obtained with switching costs is only a small proportion of the one provided when alterations are costless.

While these findings are in line with the general conjecture that switching costs are an obstacle for alterations (see Sect. 46.1), the results *do not universally support* the hypothesis that switching costs are detrimental with respect to overall performance. On the contrary, it appears that switching costs may have their beneficial effects – subject to task complexity and turbulence.

In particular, for the *decomposable* structure in Fig. 46.2a, the levels of (final and average) performance for the scenarios with costless switching is about 3 to 5.5 points of percentage higher than the respective performance levels with costly alterations. In contrast, for the *non-decomposable* structure, the performance levels achieved with high switching costs go by about 10 to 13 points of percentage beyond

[4]With $M = 4$ units of which each can switch, at maximum, two single binary choices, the maximum capacity [29] of alterations per period of the overall configuration equals 8 bits. For an indicator that is not affected by the number M of units employed in an organization, the relative average effort is normalized to the maximum number of changeable single choices per period and averaged over the observation time.

Fig. 46.2 Adaptive walks for (**a**) decomposable and (**b**) non-decomposable interaction structures for zero vs. high switching costs and no vs. moderate turbulence of the task environment. Each curve represents the average of 500 simulation runs, each run on different performance landscapes. For parameter settings see Table 46.1



the level obtained with costless alterations.[5] Moreover, it is worth mentioning that in situations with switching costs in combination with environmental turbulence,

[5]Additional experiments indicate that even when observing the organization over $T = 600$ periods (when the zero-cost/no shock scenario has come to a steady state), the performance excess obtained with high vs. zero switching costs is at a level of 7.5 points of percentage.

further performance enhancements may appear (Fig. 46.2b). Hence, given the low rates of alterations and, in particular, of false-positive moves in these scenarios (Table 46.2), this suggests that the adaptive walks with high switching costs are remarkably more efficient than with costless alterations. Regarding these dynamics two questions arise, which are discussed more into detail subsequently: (1) What causes beneficial effects of switching costs, notably when the task is not decomposable? and (2) How does task turbulence in conjunction with switching costs affect the performance levels obtained?

46.4.2 Sensitivity to Complexity and Turbulence

In the baseline scenarios, structures with a minimum and a rather high level of cross-unit interactions were simulated; in the sensitivity analysis intermediate levels of cross-unit interactions are simulated ($K^{ex} = 1$ and $K^{ex} = 3$). Moreover, the baseline scenarios capture a situation with, in fact, environmental stability and a rather moderate level of turbulence; in the sensitivity analysis, turbulence is increased stepwise, i.e., the correlation θ with the "previous" environment is reduced by 0.25 to no correlation at all ($\theta = 0$) (see Table 46.1). Figure 46.3 plots, for these different levels of cross-unit interactions and turbulence, the average performance $\overline{V}_{(0;400)}$ obtained over the observation period, differentiated for (a) costless alterations and (b) high switching costs.

The results suggest, first, that for costless alterations as well as for high switching costs, the average performances achieved for a given level of task complexity show

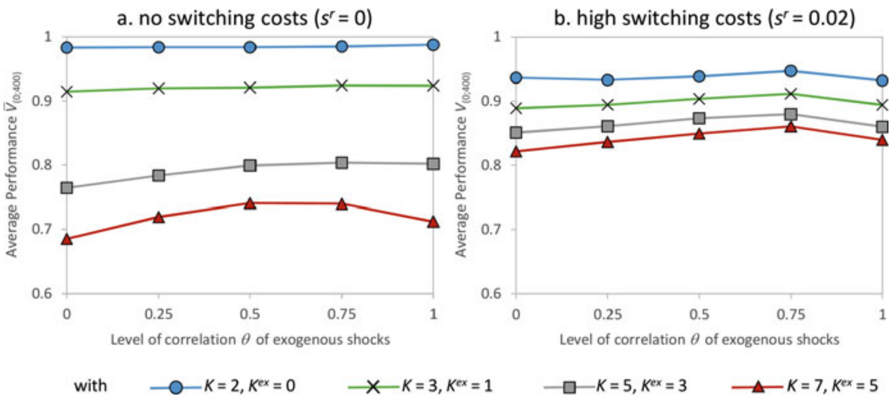


Fig. 46.3 Sensitivity of average performance $\overline{V}_{(0;400)}$ achieved over the observation time to the level of environmental turbulence θ for different levels of task complexity and for (**a**) zero vs. (**b**) high switching costs. Each mark on a curve represents the average of 500 simulation runs, each run on different performance landscapes. Confidence intervals at a level of 0.99 range from 0.0011 up to 0.013 with the highest values in all structures occurring in stable environments ($\theta = 1$). For parameter settings see Table 46.1

relatively low sensitivity to the level of environmental turbulence. However, an analysis more into detail reveals that for the vast majority of cases, the average performance achieved with uncorrelated shocks ($\theta = 0$) is lower than obtained with correlated shocks or for a stable task environment. Moreover, for high switching costs, for all levels of task complexity, not stable ($\theta = 1$) but turbulent environments undergoing highly correlated shocks ($\theta = 0.75$) show the highest average performance. For the perfectly and the nearly decomposable interaction structures (i.e., $K^{ex} \in \{0, 1\}$), switching costs lead to lower average performance than achieved with costless alterations. In contrast, for high levels of cross-unit complexity (i.e., $K^{ex} \in \{3, 5\}$), the average performance achieved with costly alterations goes remarkably beyond the level obtained with costless alterations – regardless of the level of turbulence (with the performance excess for the more complex interaction structure being even higher).

To provide an explanation for these results, the beneficial effects of switching costs in case of higher levels of cross-unit interactions serve as starting point (see question 1 in the end of Sect. 46.4.1): We argue that switching costs subtly interfere with the imperfect information that the units employ in decision-making. These imperfections result from two sources, (1) imperfect ex ante evaluations of options (see Eq. 46.7 vs. 46.9) and (2) imperfect knowledge about the fellow units' intentions. With respect to the latter aspect, recall that, in time step t, unit r forms its preferences expecting that units $q \neq r$ will stay with their choices from period $t - 1$ (Sect. 46.2.3). With this, in the end of t, unit r may not only be surprised (due to distorted ex ante evaluations) by the actual performance P^r achieved in t but also by the other units' choices, which, in case of cross-unit interactions, affect unit r's performance too. In consequence, these "surprises" may eventually let unit r adapt the "own" choices in $t + 1$, which will lead to "surprises" for units $q \neq r$ and so forth.

Hence, the "errors" in expectations due to imperfect forecasting of performance contributions and the fellow units' choices reasonably result in *frequent time-delayed mutual adjustments*. However, whether a unit leaves the status quo reasonably is shaped by the switching costs and, in particular, switching costs cause that only worthwhile alterations are made, i.e., alterations whose *perceived* performance gains exceed the costs of effort. This explanation is broadly supported by the ratio of false-positive alterations reported for the baseline scenarios in Table 46.2 which is "dramatically" reduced with switching costs. To put it in other words: *Switching costs apparently unfold a stabilizing effect in that fewer and only worthwhile alterations are implemented which obviously is the more relevant the higher the level of cross-unit interactions.*

These considerations may also provide the basis for explaining how task turbulence combined with complexity affects the performance achieved with respect to switching costs (question 2 mentioned in the end of Sect. 46.4.1): When a shock occurs, an organization is set back again to the start of search for superior solutions, i.e., to that phases of search which typically show high performance gains and frequent alterations (e.g., [13, 23]). Hence, with turbulence more alterations are likely to occur which is supported, for example, by the results displayed in Table 46.2. In this sense, turbulence mitigates the peril of inertia which is particularly relevant for high levels of complexity (as being a source of inertia by its own [1]) and may explain the performance gains induced by turbulence in the non-decomposable structure with costless alterations in Fig. 46.2b.

However, the stabilizing effect of switching costs reasonably is relevant in the early stages of search in a new environment: then performance gains are more likely to be achieved and, hence, units are more likely to alter their partial solutions. In consequence, particularly when complexity is high, units' choices induce more frequent "surprises" and time-delayed mutual adjustments as argued above. In later stages of search (when usually lower performance gains are feasible), the detrimental effects of switching costs (i.e., in the extreme, inducing inertia) predominate the benefit of increasing the efficiency of search. Obviously, with frequent shocks and, thus, an organization set back to the early stages of search again and again, these inertia-increasing effect of switching costs is less relevant and the stabilizing effects predominate.

46.5 Conclusion

The results of this simulation study suggest that the effects of switching costs subtly vary with the complexity and the turbulence of the task environment. In particular, the analysis of the simulations leads to the hypothesis that switching costs may *increase* efficiency of search for superior solutions particularly in turbulent and complex task environments. Apparently, switching costs could provide a beneficial "hurdle" preventing units within an organization from too many mutual adjustments caused by imperfect expectations on each others' decision-making behavior. This is particularly relevant when units operate in a complex and turbulent environment.

These conjectures may contribute to shedding some new light on switching costs in organizations and the discussion on obstacles for change and organizational inertia. In particular, the findings indicate on some positive effects of switching costs in organizations which is particularly relevant when switching costs are to be regarded as exogenously given, for example, by human nature, economic or technological constraints, or cultural context. With respect to the latter, the cultural dimension of uncertainty avoidance, i.e., the level to which culture programs its members to feel uncomfortable or comfortable in unstructured situations according to Hofstede [30], may induce mental switching costs. However, switching costs may also be subject to deliberate organizational design in order to stabilize the search processes of decentralized units. For example, establishing more extensive administrative procedures before a change is permitted could be regarded as a means to increase switching costs.

The simulation study calls for further research efforts, for example, to study the effects of switching costs in the context of more sophisticated coordination mechanisms than employed in this simulation model. Additionally, the interference with imperfect information on the site of decision-making agents within the organizations is to be analyzed more into detail: The limited capabilities of forecasting other agents' actions as well as the limited search space for alternative actions could be subject of promising model extensions. Moreover, a natural model extension would be to incorporate some agents' heterogeneity. For example, the units of an organization may be differently cost-efficient and may show different learning curve effects; this might result in imbalances in the effort provided to the overall organizational task as well as interesting spillover effects, particularly when the task environment is complex and turbulent.

References

1. Hannan, M.T., Freeman, J.: Structural inertia and organizational change. Am. Sociol. Rev. **49**(2), 149–164 (1984)
2. Blatt, J.M.: Optimal control with a cost of switching control. J. Aust. Math. Soc. Ser. B Appl. Math. **19**(3), 316–332 (1976)
3. Benito, G.R.G., Pedersen, T., Petersen, B.: Foreign operation methods and switching costs: conceptual issues and possible effects. Scand. J. Manag. **15**(2), 213–229 (1999)
4. Hannan, M.T., Pólos, L., Carroll, G.R.: Cascading organizational change. Org. Sci. **14** (5), 463–482 (2003)
5. Kim, H.W.: The effects of switching costs on user resistance to enterprise systems implementation. IEEE Trans. Eng. Manag. **58**(3), 471–482 (2011)
6. Polites, G.L., Karahanna, E.: Shackled to the status quo: the inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. MIS Q. **36**(1), 21–42 (2012)
7. Basaure, A., Suomi, H., Hämmäinen, H.: Transaction vs. switching costs: comparison of three core mechanisms for mobile markets. Telecommun. Policy **40**(6), 545–566 (2016)
8. Porter, M.E.: Competitive Strategy: Techniques for Analyzing Industries and Competitors. Free Press, New York (1980)
9. Burnham, T.A., Frels, J.K., Mahajan, V.: Consumer switching costs: a typology, antecedents, and consequences. J. Acad. Mark. Sci. **31**(2), 109–126 (2003)
10. Jones, M.A., Reynolds, K.E., Mothersbaugh, D.L., Beatty, S.E.: The positive and negative effects of switching costs on relational outcomes. J. Serv. Res. **9**(4), 335–355 (2007)
11. Freeman, J., Hannan, M.T.: Niche width and the dynamics of organizational populations. Am. J. Sociol. **88**(6), 1116–1145 (1983)
12. Gresov, C., Haveman, H.A., Oliva, T.A.: Organizational design, inertia and the dynamics of competitive response. Org. Sci. **4**(2), 181–208 (1993)
13. Siggelkow, N., Rivkin, J.W.: Speed and search: designing organizations for turbulence and complexity. Org. Sci. **16**(2), 101–122 (2005)
14. Chung-An, C.: Revisiting organizational age, inertia, and adaptability: developing and testing a multi-stage model in the nonprofit sector. J. Organ. Chang. Manag. **27**(2), 251–272 (2014)
15. Chang, M.-H., Harrington, J.E.: Agent-based models of organizations. In: Tesfatsion, L., Judd, K.L. (eds.) Handbook of Computational Economics: Agent-Based Computational Economics, vol. 2, pp. 1273–1337. Elsevier, Amsterdam (2006)
16. Kauffman, S.A., Levin S.: Towards a general theory of adaptive walks on rugged landscapes. J. Theor. Biol. **128**(1), 11–45 (1993)
17. Kauffman, S.A.: The Origins of Order: Self-Organization and Selection in Evolution. Oxford University Press, Oxford (1993)
18. Wall, F.: Agent-based modeling in managerial science: an illustrative survey and study. Rev. Manag. Sci. **10**(1), 135–193 (2016)

19. Li, R., Emmerich, M.M., Eggermont, J., Bovenkamp, E.P., Bäck, T., Dijkstra, J., Reiber, J.C.: Mixed-integer NK landscapes. In: Runarsson, T., Beyer, H.-G., Burke, E., Merelo-GuervÃşs, J., Whitley, L.D., Yao, X. (eds.) Parallel Problem Solving from Nature IX, LNCS, vol. 4193, pp. 42–51. Springer, Heidelberg (2006)
20. Baker, G.P.: Incentive contracts and performance measurement. J. Polit. Econ. **100**(3), 598–614 (1992)
21. Lambert, R.A.: Contracting theory and accounting. J. Account. Econ. **32**(1–3), 3–87 (2001)
22. Wall, F.: The (Beneficial) role of informational imperfections in enhancing organisational performance. In: Li Calzi, M., Milone, L., Pellizzari, P. (eds.) Progress in Artificial Economics. Lecture Notes in Economics and Mathematical Systems, vol. 645, pp. 115–126. Springer, Berlin/Heidelberg (2010)
23. Levitan, B., Kauffman, S.A.: Adaptive walks with noisy fitness measurements. Mol. Divers. **1**(1), 53–68 (1995)
24. Lorscheid, I., Heine, B.-O., Meyer, M.: Opening the “black box” of simulations: increased transparency and effective communication through the systematic design of experiments. Comput. Math. Organ. Theory **18**(1), 22–62 (2012)
25. Wall, F.: Effects of switching costs in distributed problem-solving systems. In: De la Prieta, F., Omatu, S., Fernãndez-Caballero, A. (eds.) Distributed Computing and Artificial Intelligence, 15th International Conference. Advances in Intelligent Systems and Computing, vol. 800, pp. 1–9. Springer (2019)
26. Rivkin, J.W., Siggelkow, N.: Patterned interactions in complex systems: implications for exploration. Manag. Sci. **53**(7), 1068–1085 (2007)
27. Galbraith, J.R.: Organization design: an information processing view. Interfaces **4**(3), 28–36 (1974)
28. Thompson, J.D.: Organizations in Action. Social Science Bases of Administrative Theory. McGraw-Hill, New York (1967)
29. Wall, F.: Emergence of task formation in organizations: balancing units’ competence and capacity. J. Artif. Soc. Soc. Simul. **21**(2), (2018). https://doi.org/10.18564/jasss.3679
30. Hofstede, G.J.: Culture’s Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations. Sage Publications, Thousand Oaks (2003)

Chapter 47
Governing the Digital Society: Challenges for Agent-Based Modelling

Johannes Weyer, Fabian Adelt, Sebastian Hoffmann, Julius Konrad, and Kay Cepera

Abstract Modern societies of the twenty-first century are facing a fast-growing digitisation of almost every sphere of life and work. Smart devices are collecting large amounts of data that allow to identify persons and objects and to check their current status and position. Hence, the digital real-time society is in need of new governance approaches in order to cope with these challenges.

Understanding – and maybe shaping – the digital society requires modelling the mechanisms that are guiding the dynamics of sociotechnical systems in general and of the digital society in particular. Hence, we propose a multilevel model of governance, rooted in a basic model of a sociotechnical system. This approach allows to study and to analyse complex, non-linear interactions of systems within systems and thus may provide insights how to govern the digital society and to avoid the risk of losing control.

Investigating the governance of complex sociotechnical systems requires new, innovative methods. Agent-based modelling is a suitable means to implement artificial societies at the computer screen, to experiment with various governance scenarios and to predict, which scenario produces results that are societally acceptable or politically desirable. The simulation framework SimCo, developed at TU Dortmund University and rooted in analytical sociology, helps to better understand the functioning of complex systems and to provide means for governing the digital society.

Keywords Multilevel governance · Big data · Agent-based modelling · Real-time society · Digital transformation

J. Weyer (✉) · F. Adelt · S. Hoffmann · J. Konrad · K. Cepera
Technology Studies Group, TU Dortmund University, Dortmund, Germany
e-mail: johannes.weyer@tu-dortmund.de

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_47

47.1 Digital Society

Modern societies of the twenty-first century are facing a fast-growing digitisation of almost every sphere of life and work, from industrial production to transport logistics, retailing and, finally, private life such as sleeping, cooking, leisure activities, amongst others. Smart devices are collecting large amounts of data that allow to identify persons and objects and to check their current status and position. It rather seems that algorithms are governing our world [18].

Hence, the digital society is in need of new governance approaches in order to cope with these challenges, to keep pace with the rapid technological development, to maintain control and to guarantee a worthwhile future of our society. Our main research question is how complex systems can be governed and which mode of governance (soft vs. strong control) produces best results. As shown later in Sect. 47.4, agent-based modelling with the simulation framework SimCo, developed at TU Dortmund University, may help to better understand the functioning of complex systems and to provide means to govern the digital society.

47.1.1 Process Model

In order to understand what happens with our data, we developed a process model with three stages: data acquisition, data analysis and use of data (see Fig. 47.1).

In the first step, data are collected by humans (e.g. via smart watches or smartphones) or by machines (e.g. driver assistance systems in cars) that are

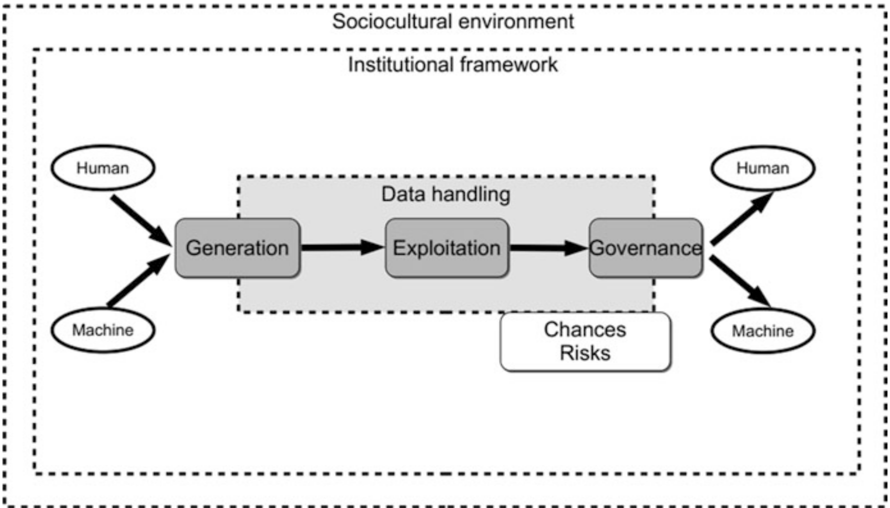


Fig. 47.1 Process model. (Source: [33]: 74)

transmitted to various data analysts [33]. In the second step, these data are processed in order to better understand the individual, her/his preferences, habits and desires, for example, for marketing purposes. Apart from this targeting single individuals, data can also be used to generate a situation picture about the current state of the whole system (e.g. road transportation), where the individual is only a small part, whose properties are less relevant [14].

47.1.2 Data Use and Misuse

Most debates about data society stop here and discusss issues of privacy and data protection [10] which, doubtless, are very important and have to be taken seriously. However, abuse of data is only one side of the coin; the other side is the legal use of data analysis for controlling people and machines. This is the third step in our model, pointing at the gigantic potential of data analysis for controlling individual people, for example, by giving advices for healthier nutrition, or even whole systems such as the transportation or the energy system. As Sect. 47.2 will show in more detail, complex systems that seem to be uncontrollable nowadays are controlled by algorithms, navigating our cars or switching on and off our solar panels.

47.1.3 Trust and Institutions

Governance by algorithms thus has become part of life and work in the digital society [13]. However, this new kind of governance only works if two conditions are met: all partners should be trustworthy, and an institutional, legal framework should protect transactions. People, willing to transmit data (in Step 1), have to trust data analysts; otherwise, they will refrain from using their services, for example, following their advices in Step 3 [3]. Vice versa, data analysts have to be sure that transmitted data are not manipulated [19]. Additionally, similar to life in the analogue world, an institutional framework is needed to protect all participants from fraud, abuse or other illegal activities. Hence, the state has to redefine its role in digital societies, setting the boundary conditions for digital transactions and interactions that are hardly controllable in detail [8].

47.1.4 Individual Impacts of Digital Transformation

Sociology mostly has put emphasis on the impacts of digital transformation regarding the individual. Apart from data protection concerns, the issues of acceleration and alienation have been raised by Sherry Turkle [27], Hartmut Rosa [23, 25] and others.

Turkle radically distinguishes between the real world of 'conversation' and the virtual world of – technically mediated – 'connection' [28: 1], where we lose our ability to communicate and finally are 'alone together' (title of her book). Computerisation and digitisation are regarded as solely negative, leading to a loss of humanity and eroding interpersonal relationships.

The same applies to Rosa, who doubts that 'good, self-determined life' still is possible, if every societal process is permanently accelerated [24]. We increase the number of actions per period of time, but finally we fail and get exhausted and burned out. Although nobody would doubt his diagnosis to be at least partly true, it doesn't cover the whole picture. First, he (and Turkle as well) only puts emphasis on the individual overload (as psychologists typically would do); second, he rates all these developments as negative (not revealing his valuation standard except from the concept of 'good life', taken from philosophy); and third, he hardly reflects the societal consequences of digitisation and acceleration, as sociologist typically would do.

47.1.5 Societal Impacts of Digital Transformation

Hence, we propose to take a closer look at the societal dimension, reflecting the dynamics and the changing structure of complex sociotechnical systems as well as the impact of these transformation processes on politics and the state, on various societal subsystems and, finally, on the individual. The main issue will be: Will the digital society be governed by autonomous technological systems and stubborn algorithms and finally get out of control, or will we find measures to cope with the challenges of governing the digital society? As a first step towards governance of complex systems, we are in need of understanding their functioning.

47.2 Real-Time Society

Starting with the diagnosis of Rosa, we can observe an acceleration of almost every societal process, triggered by digitisation. Compared with a letter, that took days or even weeks, if delivered abroad, the email can be (or even has to be) answered immediately nowadays. Suppliers deliver their products just in time to the production line, controlled by OEM's smart enterprise resource planning system. Furthermore, route guidance systems deliver advices almost in real time, dispensing the necessity to plan routes in advance by means of printed maps.

The digital society has become a real-time society, since smart, networked devices allow to permanently update the data base as well as the situation picture and to communicate relevant information to all participants – again almost in real-time. Actors adjust their plans and subsequent actions according to the current system's state at point of time t_x, thus changing this state and forcing the system to adapt itself – and to deliver new recommendations, based on the system's state at point of time $t_x + 1$.

We propose – by taking Rosa's argument one step further – that this feedback loop, processed in real-time, is at the core of acceleration or, as we would call it, of the real-time society, where processes of planning and acting no longer take place sequentially (with large time slots in between), but rather simultaneously. However, we refrain from jumping to a negative evaluation of these developments, but call for an unbiased empirical investigation of the real-time society, its dynamics and its governance [30].

47.2.1 Practical Examples

In practise, we can observe various examples of real-time control of complex systems, for example, in industrial production and logistics, in road transportation or in concepts of future smart grids. The basic mechanisms of central control of decentralised systems have even been described 20 years ago by Gene Rochlin [22] – again negatively with the phrase 'trapped in the net' (title of his book). He argued that the new options will not serve as a means of liberation and democratisation, but mainly as a means of control, since data networks allow for a previously unknown micromanagement of decentralised processes, for example, decisions or actions of individuals, by a central controller.

Looking at the example of route guidance systems for road transportation such as TomTom, Google Maps or others at first glance seems to confirm Rochlin's diagnosis. A central controller collects large amounts of data from users, computes a situational picture and issues recommendations in real time. However, the users are free to follow those recommendations or to ignore them. If they are able to plan independently (e.g. because of specific knowledge about the local situation), they can make their own strategic choices. If not, they may become increasingly dependent on the system, forcing them to adapt to the respective situation passively. The system's state thus emerges – and dynamically develops – as a result of the interplay of various independent decisions of decentralised decision makers, who are influenced by interventions of the central controller.

47.2.2 New Mode of Governance?

This smart mixture of central control and decentralised self-control has emerged as a new pattern or new mode of governance in practise, but has hardly been debated or even understood by governance research [11]. We suppose that the digital society will need to better understand these mechanisms in order to maintain control and to develop options of smart governance. However, the political regulation of algorithms [26] presupposes a deeper knowledge of the dynamics of complex systems and their governance. As mentioned before, ABM may help to gain insights into this topic.

47.3 Modelling of Complex Systems

In order to implement the governance of digital societies by means of agent-based modelling (ABM), we first need a conceptual model of the governance of complex sociotechnical systems, which will be developed very briefly in this section.

47.3.1 Basic Model of a Sociotechnical System

Our starting point is a model of a social or sociotechnical system that contains the components (actors), the rules of (inter-)action and the mechanism of structure formation, which are presented in a formal, theory-based language [12]. We refer to the macro-micro-macro model of James Coleman [4], the Institutional Analysis and Development (IAD) framework of Elinor Ostrom [20, 21] and the model of social explanation (MSE) of Hartmut Esser [5].

This basic model, in which actors act and interact uncoordinatedly and the system structure emerges spontaneously from their interactions, will serve as starting point for more elaborated extensions, applying this model to the study of governance of complex systems.

47.3.2 Multilevel Governance

Referring to a long-lasting debate on governance theory [11], we propose to use the term ‘governance’ for ‘a specific combination of the basic mechanisms control and coordination in multi-level socio technical systems’ [31: 17]. We assume that these two basic mechanisms rule the internal operations of social or sociotechnical systems. Additionally, they are used to influence other systems, which on their part are guided by an internal mechanism and – with regard to their openness or closeness – can react differently to external attempts of control or coordination [17].

Control is defined as ‘the intentional intervention into socio-technical systems, aiming at producing intended effects’ [31: 18], whereas coordination is ‘the mutual adjustment of heterogeneous actors aiming at collectively solving problems in a way that is acceptable to all parties involved’ (ibid.: 22).

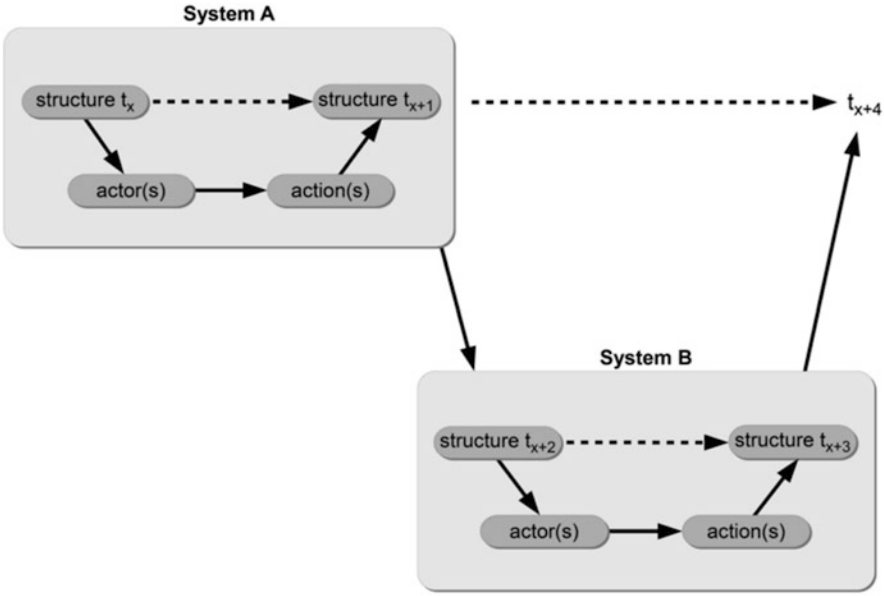


Fig. 47.2 Control in a multilevel configuration. Source: [31: 22]

As depicted in Fig. 47.2, control is an attempt of external actors (in system A), to purposely influence the state and the processes of system B. In an action theoretical perspective, control is mostly exerted indirectly by changing the situational parameters of system B at point of time t_x, which are part of the actors' decision-making and thus enabling the controller to achieve the presumed outcome at point of time $t_x + 1$ via this 'detour'.

However, it is hard to predict (and even more difficult to control), whether system B produces the desired outcome or rather unintended effects. Hence, any attempt to control other systems (or actors within other system) faces the risk of failure.

One can zoom in into the respective systems and observe the internal mechanisms, which may be control or coordination – or even a mixture of both. Or one can zoom out and observe the interrelations and interactions of various systems, which mutually influence each other by inputs (downward arrow) or feedbacks (upward arrow).[1]

Further extending this idea, we can conceive of governance as multilevel governance – a multilayer structure, where various systems interact and mutually influence each other. As mentioned above, it is a nested structure of systems of systems, which can be decomposed into even smaller units or taken as a whole.

[1] Figure 47.2 does not presuppose that relations have to be hierarchical.

47.4 Challenges for ABM

At first glance, a multilevel structure seems to be too complicated for sociological analysis, since it may entail a large number of systems and interactions. However, ABM may help to model and to understand the dynamics and the governance of those nested structures, sticking to a few basic software modules that can be utilised throughout the model, namely, a basic model of a sociotechnical system, a decision algorithm and two basic interaction mechanisms (control and coordination, as depicted above).

The central principle of constructing virtual images of complex real-life systems is to link various systems by taking the output of system A as input to system B and vice versa, as depicted in Fig. 47.2. System interaction thus means setting the boundary conditions of other systems and to receive their feedback as input to the respective system. This way, even very complex, multilevel structures can be constructed by using a manageable number of software modules. Furthermore, the results of non-linear interactions can be observed and analysed.

47.4.1 Algorithmic Actions

Human actors and software agents may be part of this virtual society. We assume that most algorithms (e.g. driver assistance systems) can be technically implemented by a utility-maximising procedure, similar to the one we use for agents representing humans [9]. Hence, the main challenge is not to implement smart technology, but to make models of human actors realistic by taking into account real-life decisions of heterogeneous actors and various actor types. Modelling human actors as software agents thus allows studying the interaction of (artificial) humans and algorithms (representing smart technology).

47.4.2 System Control

Furthermore, this ABM approach lets us control different parameters and study various scenarios of the digital society, for example, by changing the degree of autonomy of smart machines. Additionally, we can observe the effects of interventions by measuring various outcomes and comparing results of different settings or scenarios. This makes it easier to answer questions about the future of the digital society, about the effects of real-time operations and, finally, about the political regulation of complex, self-organised sociotechnical systems.

47.4.3 SimCo

In the years, we started to work on multilevel governance by means of ABM at TU Dortmund University. We developed the simulator SimCo ('Simulation of the Governance of Complex Systems') and conducted several experiments on risk management and system transformation, mostly in road transportation [2, 32]. SimCo is a first step to cover the whole picture of the multilevel architecture. Most work until now has put emphasis on the operational processes of urban road transportation, that is, the daily decisions on technology and route choice.

SimCo is rooted in a sociological macro-micro-macro model of a sociotechnical system, referring to sociological theories of behaviour and decision-making [6, 7]. Governance measures do not necessarily have a direct impact on the structures of a sociotechnical system, but rather influence decision making of actors at the micro level which then leads to emergent effects on the macro level (e.g. a regime shift, lowering congestion).

SimCo has been developed as a general-purpose framework. It contains dimensions that are freely programmable, allowing us to conduct experiments with various scenarios, for example, shifting road transportation towards sustainability.

In this scenario, agents' transport mode choices are important. They are shaped by their individual perception of situational constraints and their individual preferences. Situational constraints may change due to the system's dynamics (e.g. in case of congestion) or due to policy interventions (e.g. in the case of banning diesel cars).

Agents take decisions with bounded rationality and under uncertainty, referring to multiple evaluation criteria [29]. Behavioural alternatives are evaluated by assigning a utility to every possible consequence, and agents finally choose the option with the highest subjective expected utility [15].

47.4.4 Empirical Results

We conducted several simulation experiments with various what-if scenarios and tested different governance measures regarding their efficiency and unintended side effects (cf. [2]).

One major result of our experiments is: Political interventions, for example, to minimise risks (congestions or emissions) or to change the system towards sustainability work best, if the governance mode of soft control is applied (cf. Fig. 47.3), working with incentives and not with strong measures such as bans (as in the case of strong control) (cf. also [1]).

Experiments on real-time governance have also been conducted, comparing the traditional mode of sequential planning and action with the real-time, simultaneous mode, where plans are no longer made in advance, but decisions are framed and guided by real-time information (e.g. route guidance in road transportation) [16]. We

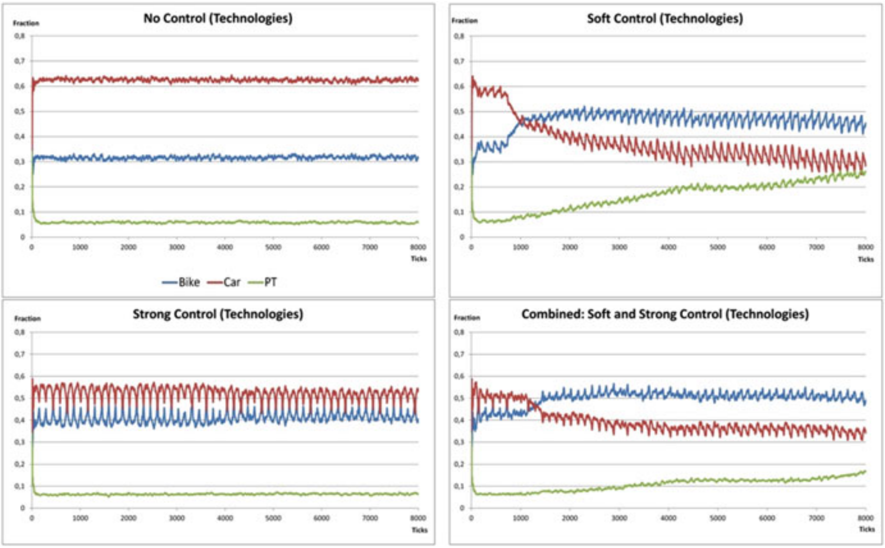


Fig. 47.3 Technology usage (fraction) in course of time (8000 ticks)

have analysed the effects of real-time governance at the system level (by measuring the efficiency of the system) and at the actor level (by measuring the satisfaction of actors).

47.5 Conclusion

Understanding the challenges of the digital society requires understanding the mechanisms that are guiding the dynamics of sociotechnical systems in general and of the digital society in particular. We propose a multilevel model of governance, grounded in a basic model of a sociotechnical system that allows to study and to analyse complex, nonlinear interactions of systems within systems and thus may provide insights how to govern the digital society, inhabited by smart humans and smart machines, and to avoid the risk of losing control. Political regulation will no longer rely on traditional measures of the past, but has to become 'smart' as well [34]. ABM allows to experiment with various types of interventions and to find out which one works best and helps to shape a human future of the digital society.

References

1. Adelt, F., Weyer, J., Fink, R.D.: Governance of complex systems. Results of a sociological simulation experiment. Ergonomics (Special Issue "Beyond Human-Centered Automation"). **57**, 434–448 (2014). https://doi.org/10.1080/00140139.2013.877598
2. Adelt, F., Weyer, J., Hoffmann, S., Ihrig, A.: Simulation of the governance of complex systems (SimCo). Basic concepts and experiments on urban transportation. J. Artif. Soc. Soc. Simul. **21**(2), (2018). http://jasss.soc.surrey.ac.uk/21/2/2.html
3. Cepera, K.P., Weyer, J., Konrad, J.: App-solut! Vertrauen in Mobile Applikationen (Apps) (Soziologische Arbeitspapiere 58/2019). TU Dortmund, Dortmund (2019)
4. Coleman, J.S.: Foundations of Social Theory. Harvard University Press, Cambridge, MA (1990)
5. Esser, H.: The rationality of everyday behavior: a rational choice reconstruction of the theory of action by Alfred Schutz. Rational. Soc. **5**, 7–31 (1993)
6. Esser, H.: Soziologie. Allgemeine Grundlagen. Campus, Frankfurt/M (1993)
7. Esser, H.: Soziologie. Spezielle Grundlagen, Bd. 5: Institutionen. Campus, Frankfurt/M (2000)
8. Feick, J., Werle, R.: Regulation of cyberspace. In: Baldwin, R., Cave, M., Lodge, M. (eds.) Regulation, pp. 523–547. Oxford University Press, Oxford (2010)
9. Fink, R.D., Weyer, J.: Interaction of human actors and non-human agents. A sociological simulation model of hybrid systems. Sci. Technol. Innov. Stud. **10**, 31–46 (2014)., www.sti-studies.de/ojs/index.php/sti/article/view/131/127
10. Genner, S.: ON/OFF: Risks and Rewards of the Anytime-Anywhere Internet. vdf Hochschulverlag AG (2017)
11. Grande, E.: Governance-Forschung in der Governance-Falle?–Eine kritische Bestandsaufnahme. Polit. Vierteljahresschr. **53**(4), 565–592 (2012)
12. Hedström, P., Swedberg, R.: Social mechanisms. Acta. Sociol. **39**, 281–308 (1996)
13. Just, N., Latzer, M.: Governance by algorithms: reality construction by algorithmic selection on the internet. Media. Cult. Soc. **39**, 238–258 (2016). http://journals.sagepub.com/doi/pdf/10.1177/0163443716643157
14. Kappler, K., Schrape, J.-F., Ulbricht, L., Weyer, J.: Societal implications of big data. KI Künstliche Intelligenz, Special Issue "Big Data". **32**(1), 55–60 (2018). https://doi.org/10.1007/s13218-017-0520-x
15. Konidari, P., Mavrakis, D.: A multi-criteria evaluation method for climate change mitigation policy instruments. Energy Policy. **35**(12), 6235–6257 (2007)
16. Konrad, J., Weyer, J., Cepera, K., Adelt, F.: Echtzeitsteuerung komplexer Systeme. Eine Simulationsstudie (Soziologische Arbeitspapiere 57/2019). TU Dortmund, Dortmund (2019)
17. Luhmann, N.: Die Wissenschaft der Gesellschaft. Suhrkamp, Frankfurt/M (1990)
18. O'Reilly, T.: Open data and algorithmic regulation. In: Goldstein, B., Dyson, L. (eds.) Beyond Transparency: Open Data and the Future of Civic Innovation, pp. 289–300. Code for America Press, San Francisco (2013)
19. Open Effect: Every Step You Fake. A Comparative Analysis of Fitness Tracker Privacy and Security. In, https://openeffect.ca/reports/Every_Step_You_Fake.pdf (2016)
20. Ostrom, E.: Understanding Institutional Diversity. Princeton University Press, Princeton (2005)
21. Ostrom, E.: Beyond markets and states: polycentric governance of complex economic systems. Am. Econ. Rev. **100**, 641–672 (2010)
22. Rochlin, G.I.: Trapped in the Net. The Unanticipated Consequences of Computerization. Princeton UP, Princeton (1997)
23. Rosa, H.: Beschleunigung. Die Veränderung der Zeitstrukturen in der Moderne. Suhrkamp, Frankfurt/M (2005)
24. Rosa, H.: Interview mit SZ Wissen. In, http://www.eilkrankheit.de/Interviews/iv2.pdf (2008)
25. Rosa, H.: Weltbeziehungen im Zeitalter der Beschleunigung: Umrisse einer neuen Gesellschaftskritik. Suhrkamp Verlag, Frankfurt/M (2012)

26. Saurwein, F., Just, N., Latzer, M.: Governance of algorithms: options and limitations. Info. **17**(6), 35–49 (2015). https://doi.org/10.1108/info-05-2015-0025
27. Turkle, S.: Alone Together. Why We Expect More from Technology and Less from Each Other. Basic Books, New York (2011)
28. Turkle, S.: The flight from conversation. In: The New York Times, pp. 22, http://www.nytimes.com/2012/04/22/opinion/sunday/the-flight-from-conversation.html (2012)
29. Velasquez, M., Hester, P.T.: An analysis of multi-criteria decision making methods. Int. J. Operat. Res. **10**(2), 56–66 (2013)
30. Weyer, J.: Die Echtzeitgesellschaft. Wie smarte Technik unser Leben steuert. Campus, Frankfurt/M (2019)
31. Weyer, J., Adelt, F., Hoffmann, S.: Governance of complex systems. In: A Multi-level Model (Soziologisches Arbeitspapier 42/2015). Dortmund, TU Dortmund (2015). http://hdl.handle.net/2003/34132
32. Weyer, J., Adelt, F., Hoffmann, S.: Governance of transitions. A simulation experiment on urban transportation. In: Squazzoni, F. (ed.) Social Simulation for a Digital Society: Applications and Innovations in Computational Social Science. Springer, Berlin (2019). (forthcoming)
33. Weyer, J., Delisle, M., Kappler, K., Kiehl, M., Merz, C., Schrape, J.-F.: Big data in soziologischer Perspektive. In: Hoeren, T., Kolany-Raiser, B. (eds.) Big Data und Gesellschaft. Eine multidisziplinäre Annäherung, pp. 69–149. Springer, Berlin (2018)
34. Willke, H.: Smart Governance. Governing the Global Knowledge Society. Campus, Frankfurt/M (2007)

Chapter 48
Towards Modelling Interventions in Small-Scale Fisheries

Nanda Wijermans and Elizabeth Drury O'Neill

Abstract Interventions often have unintended effects, particularly when they target outcomes in complex social-ecological systems (SES), such as fisheries. Development agencies strive for 'doing development differently', because past efforts have often been unsuccessful. The question then is how to do better. Science can make a major contribution here. Its power lies in synthesising and generalising knowledge embedded in 'thinking tools' that enable looking at problems from a complex adaptive systems lens. Our approach is to design an agent-based model to understand and reflect on the impact of different types of development interventions in small-scale fisheries. The model embeds synthesised insights from two small-scale fishery cases and from literature on important (loan and entry–exit) dynamics. In short, by modelling a complex socioecological phenomenon such as fishing and trading in small-scale fisheries, the model aims at enabling scientist and practitioners alike to think differently and thereby to support acting differently.

Keywords Social-ecological modelling · Fisher behaviour · Development

48.1 Introduction

Interventions[1] often have unintended effects, particularly when they target outcomes in complex social-ecological systems (SES), such as fisheries. Development agencies strive for 'doing development differently', because past efforts have often been unsuccessful. The question then is how to do better. Science can

[1]The word intervention is used to refer to purposeful action to change a situation for a particular goal. Who initiates or designs this action is not specified, could thus both be either an external or the target of the change.

N. Wijermans (✉) · E. Drury O'Neill
Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden
e-mail: nanda.wijermans@su.se

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_48

make a major contribution here. Its power lies in synthesising and generalising knowledge embedded in 'thinking tools' that enable looking at problems from a complex adaptive systems lens. Our approach is to design an agent-based model to understand and reflect on the impact of different types of development interventions in small-scale fisheries. The model embeds synthesised insights from two small-scale fishery cases and from literature on important (loan and entry–exit) dynamics. In short, by modelling a complex social-ecological phenomenon such as fishing and trading in small-scale fisheries, the model aims at enabling scientist and practitioners alike to think differently and thereby to support acting differently.

Small-scale fisheries (SSF) are often referred to as traditional, artisanal, low-tech, labour intensive, low capital operations that differ in particular in comparison with their large-scale opposites in terms of, for example, their social organisation, economic motivation and market linkages. SSF provide a major contribution to global fish catch (about 50%) and employment for millions of people [1]. Their sustainability is thus critical for global fish production and local poverty reduction. However, fishing is a particularly risky livelihood in comparison with other livelihoods, such as agriculture [2]. Moreover, fishers are more likely to be affected by and vulnerable to factors such as climate change, for example, by being struck by natural disasters or reduction in catch due to change in variability and predictability of the stock. The recent decline in fish catches [2–5] has been attributed to too many boats fishing, protected areas reducing fish access or environmentally destructive fishing. This will affect livelihoods or make them unviable. Overall, it is an extremely complex task to intervene in SSF. To deal with these challenges, it is essential to understand the SSF as a social-ecological system.

Social-ecological systems (SES) are characterised by having social and ecological dynamics that continuously affect each other. In many cases, interventions and actions taken by governments, NGOs and/or researchers in fishing communities seem to be working with one of two objectives: a social objective (e.g. poverty reduction) or an ecological objective (e.g. the conservation of natural resources) [3]. However, the effect of, for example, a successful ban on the fishing (ecological) may lead to a recovery of the fish stock, but may have a detrimental effect on all the livelihoods that depend on fishing (social). This is but one obvious blind spot an SES lens can address by identifying relevant actors, factors, their interactions and mechanisms.

We therefore propose to develop an ABM that aims to represent a typical small-scale fishery system for supporting thinking. To this end, the model needs to reflect the synthesis of insights from multiple small-scale fishery case studies, herein fisheries in the Philippines and Zanzibar, literature on key dynamics and development interventions. The model is thus informed by cases, rather than modelling a specific case. This extended abstract reports on the design stage of the model: from different sources of knowledge (case, literature, workshop) to the design of a conceptual model.

48.2 Synthesising Knowledge and Insights

The small-scale fishery case studies included are the Philippines (PHI) and Zanzibar (ZA) fisheries; see [6] for more information on the ZA case study. We gathered general information (geographical, sociocultural, livelihood security and ecological) and fishery specifics (characteristics, target species and market (access)) and identified main actors and interactions; see for the work in Zanzibar. The cases were contrasted to identify critical model elements, see Table 48.1.

The literature provides insights on key dynamics in SSF, in particular, concerning loans and entry–exit dynamics. Note, this concerns SSF in general, not specifically for PHI and ZA. Examples across numerous cases include noticeable differences in loan interest rates with some fisheries not charging interest and other fisheries having extremely high interest rates, for example, that newer fishers in general are more willing to exit than fishers that are fishing for >10 years [7], whilst exit choice dynamics are highly context dependent on, for example, the increase or decrease of livelihood alternatives, catch and wealth.

Development interventions and their effects hardly appeared in the reviewed literature (exception is, e.g., [8]). As an alternative to literature, we therefore collected narratives of interventions to have a benchmark. Since the second author returned to the field (PHI), she integrated focus group interviews in her. In addition, we conducted a workshop with fishery researchers and practitioners. From this set of narratives, we derived three archetypes of interventions: capacity building, post-event, and conservation, see Table 48.2.

We also used the workshop to identify potential blind spots in the narratives which can inform our model focus and choices, with particular attention paid to the contribution of an explicit SES lens on foreseeing the consequences of interventions that can help avoid unintended effects in future.

Table 48.1 Contrasting the two cases (excerpt)

Similarities	Differences
Fisheries. Tropical coral reef marine ecosystems, multi-gear and multi-species fisheries *Livelihood*. Depends on fishing activities for income, where fishers tend to earn less than traders. However, most fishers are in a trader–fisher relation, that is, an informal contract for loan and favours *Problems*. The landings are in decline in terms of weight and diversity of species composition	*Gender*. Men and women work together (PHI) versus separated (ZA). Fisherwomen earn significantly less than men (PHI), trading women earn less than men (ZA) *Livelihood*. Switching livelihoods occurs less in (PHI) than in (ZA) *Market*. Sales are mostly commercialised in PHI and centred around export. ZA however is connected to tourism *Trader–fisher relationship*. It is more prevalent and institutionalised in PHI (95%) than in ZA (55%)

Table 48.2 Intervention archetypes from the collected intervention narratives

Type of interventions	Underlying reasoning	Form
Capacity building *Social interventions, social capacity and power relations*	Supports stronger societies through *knowledge* about their rights that they can self-push a change (e.g. asking the government for better infrastructure)	Workshops
Post-event *Aid–relief after (natural) disasters*	Support fishers to overcome the damage and keep on fishing and replace damaged goods and infrastructure	Providing boats, gear Building infrastructure: processing plants
Conservation	Stock needs to recover; reduce pressure by removing fishers	Fish ban

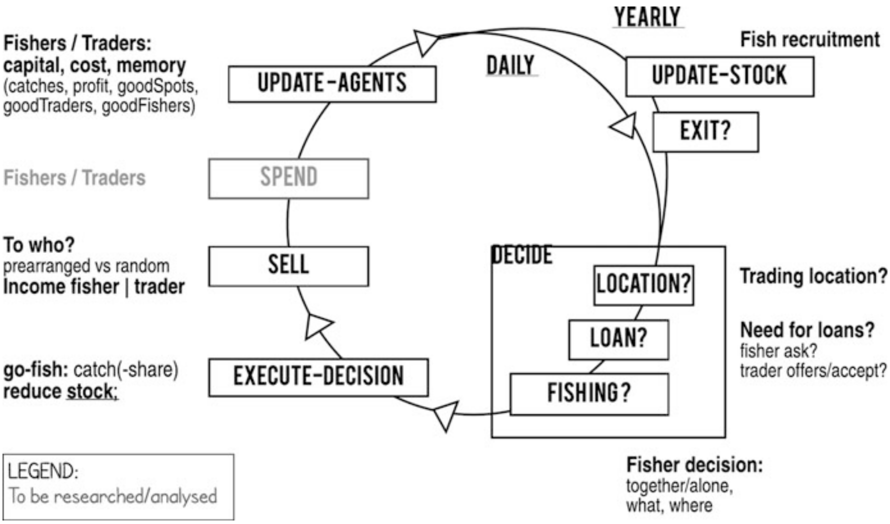


Fig. 48.1 Conceptual model of the model process

48.3 Conceptual Model (PHIZA-FISH)

The case reflections, knowledge from literature on key dynamics and the collected narratives on development interventions allowed us to make choices for the conceptual model (process), illustrated in Fig. 48.1. Every tick fishers and traders decide and execute their choices; they sell the fish and potentially spend their income. The fish stock dynamics are on another time scale and update on yearly basis.

48.4 Conclusion

The model is now for 50% construction and requires a few more iterations to start to explore the role of interventions. However, the process of designing a model reflecting the existing knowledge of multiple cases and general literature insights is something that needs to be carefully crafted, as model choices have a tremendous influence on the behaviour the model will display. The presentation at the conference thus invites feedback from modelling peers on the use of different sources of knowledge for model design and the usefulness of models as thinking tools.

References

1. Islam, M.M., Sallu, S., Hubacek, K., Paavola, J.: Limits and barriers to adaptation to climate variability and change in Bangladeshi coastal fishing communities. Mar. Policy. **43**, 208–216 (2014)
2. Crona, B., Nyström, M., Folke, C., Jiddawi, N.: Middlemen, a critical social-ecological link in coastal communities of Kenya and Zanzibar. Mar. Policy. **34**, 761–771 (2010)
3. Brooks, S.E., Kebede, B., Allison, E.H., Reynolds, J.D.: The balance of power in rural marketing networks: a case study of Snake trading in Cambodia. J. Dev. Stud. **46**, 1003–1025 (2010)
4. Haque, C.E., Idrobo, C.J., Berkes, F., Giesbrecht, D.: Small-scale fishers' adaptations to change_ the role of formal and informal credit in Paraty, Brazil. Mar. Policy. **51**, 401–407 (2015)
5. Chambers, C., Carothers, C.: Thirty years after privatization_ a survey of Icelandic small-boat fishermen. Mar. Policy. **80**(1–12), (2017)
6. O'Neill, E.D., Crona, B.: Assistance networks in seafood trade – a means to assess benefit distribution in small-scale fisheries. Mar. Policy. **78**, 196–205 (2017)
7. Muallil, R.N., Geronimo, R.C., Cleland, D., Cabral, R.B., Doctor, M.V., Cruz-Trinidad, A., Aliño, P.M.: Willingness to exit the artisanal fishery as a response to scenarios of declining catch or increasing monetary incentives. Fish. Res. **111**, 74–81 (2011)
8. Tewfik, A., Andrew, N.L., Béné, C., Garces, L.: Reconciling poverty alleviation with reduction in fisheries capacity: boat aid in post-tsunami Aceh, Indonesia. Fisheries Manage. **15**, 147–158 (2008)

Chapter 49
Population Characteristics and the Decision to Convert to Organic Farming

Q. Xu, S. Huet, E. Perret, I. Boisdon, and G. Deffuant

Abstract We revisit some ideas of why farmers do not convert to organic farming from our previous article with a dynamic individual-based model. In this model, an agent's decision on transitioning to organic is based on the comparison between the satisfaction with its current situation and the potential satisfaction with an alternative farming strategy. A farmer agent's satisfaction is modeled with the theory of reasoned action. It is computed by comparing the agent's outcomes over time and comparing its current outcome against those of other agents to whom it lends great credibility ("important others"). The first study is based on prototypical farm populations. In this chapter, the predicted conversion rate is studied with some French "cantons" having different practice intensities. The model is initialized with dairy farmers' data in these "cantons" in 2000. The results show that the "cantons" characteristics have great impact on the virtual adoption rate. Intensive "cantons" convert less than extensive ones. Extensive farms having not very good environmental outcomes seem to convert the most.

Keywords Organic farming · Decision-making · Major change · Theory of reasoned action · Agent-based model · Social influence · Credibility

Q. Xu (✉)
School of Information Engineering, Jiangxi University of Science and Technology, Ganzhou, China

UCA, AgroParisTech, Inra, Irstea, VetAgro Sup, Territoires, Paris, France
e-mail: xuqing@jxust.edu.cn

E. Perret · I. Boisdon
UCA, AgroParisTech, Inra, Irstea, VetAgro Sup, Territoires, Paris, France

S. Huet (✉) · G. Deffuant
UCA Lapsco, INRAE Lisc – INRAE, Aubière, France
e-mail: sylvie.huet@inrae.fr

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5_49

49.1 Introduction

The recent dairy crisis combined with increasing consumer demand for organic food has made conversion to organic farming socially and economically interesting for dairy farmers [1, 2]. However, many farmers still do not convert. Why?

Recently, the conversion to organic farming has been qualified as a major change [3] or a transformational adaptation [4, 5] and a social movement [3, 6–9]. Conversion often implies strong changes in a farmer agent's worldview and social network and generally begins with a strong need for change [3, 10]. Such a change that engages a number of social processes involving the agent, its peers, and its environment has been rarely studied [11].

Agent-based modeling [12] or individual-based modeling [13] appears to be relevant and well-geared means to help identify the main drivers that can explain the observed dynamics. However, as pointed out in our previous work [14, 15], none of the current agent models is well fit to represent the decision process about a major change that is at stake in the conversion to organic farming. This is the reason why we have proposed a dynamic version of the theory of reasoned action (TRA) and showed its potential to explain why farmers do not convert to organic farming in a prototypical farmer population [14]. This chapter aims at evaluating this model when running on more realistic farmer populations designed from the French agricultural census data (RGA) 2000. The purpose is mainly to study how the conversion rate evolves differently in various farmer populations. The main original features of the model are the following:

- Like Kaufmann [16], we propose a dynamic model of TRA [17] to compute an agent's satisfaction. In the model, a decision on conversion is based on comparing the agent's satisfaction with its current strategy against its potential satisfaction with an alternative one. These satisfactions are computed with the attitude and the subjective norm related to the current or alternative strategy. Both attitude and subjective norm are dynamic and based on the difference in practice outcomes. Consistently with TRA [17], our model integrates decomposed variables: attitude and subjective norm—instead of global variables as the model of Kaufmann [16]. A farmer agent's evaluation is thus based on its concrete strategy and practice instead of an abstract general opinion. The evaluation relates to its own experiences and its peers' strategies and practices.
- An agent's attitude about its current strategy is modeled as the difference between its past outcomes with current strategy and its current outcome. This choice is grounded in the propositions of Mintzberg and Öhlmér [18, 19] who argue that a farmer agent's earlier practices are very influential for its decision on a future practice. Regarding an alternative strategy which can be new to the agent, the attitude is the difference between its past outcomes with current strategy and the outcomes of other peers with a similar farm and adopting the alternative or the outcome of the alternative in the media.

- The subjective norm involves a comparison with peers, weighted by their credibility. This credibility is based on the difference between outcomes (i.e., total milk production).
- As Darré [20] showed that farmers co-construct their practices, an agent updates its practice by copying the performances of its credible peers having similar farm characteristics (number of dairy cows, utilized agricultural area, etc.).
- Finally, an important feature of our model is tied to the decision process about a "major change." It is assumed that a major change is only considered in critical situations where an agent faces high costs (economic, cognitive, emotional, etc.). Otherwise, agents do not even consider changing their major options. In a stable period, if a farmer agent is satisfied enough with its current farming strategy, it does not envisage an alternative one. Only a certain level of frustration or critical event (succession, major change in the farm, etc.) will evoke a consideration of an alternative to a farmer's current strategy.

In our previous work [14], the model is studied with prototypical farm populations and shows different reasons for the lack of conversion to organic farming. In this work, the model is explored further by an initialization with six French "cantons" (100 farmers in each "canton") varying in their distributions of farm characteristics. The adoption degree is compared among these populations to clarify the effect of population features. In general, intensive (with a high input and high output philosophy) "cantons" convert less than extensive (with a low input and low output philosophy) ones. Particularly, extensive farms having not very good environmental outcomes seem to convert the most.

After presenting the model's principles with partially ODD (overview, design concepts, details) protocol [21], we outline the model's behavior and some explanations before going on to synthesis and discuss our conclusions.

49.2 Materials and Methods

49.2.1 The Model

49.2.1.1 Basic Elements

Farmer

The model studies the evolution of a population of N farmer agents. Each agent f is characterized by its farm, its farming strategy, its practice defined on several dimensions i which are evaluated through performances, the importance W_i given to each dimension of practice, the credibility ($C(f,v)$) it lends to each other agent v, its memory of applied strategies and practices during the last M periods, its satisfactions with current farming strategy (I_C) and with an alternative one (I_A), and its duration for staying with a strategy (DC) and for being dissatisfied for a strategy after being stable (DD).

DC and DD capture the duration between two events related to the decision process. DD counts an agent's dissatisfaction duration with its current strategy. In the model, an agent has to be dissatisfied long enough with its current strategy to change it. DC counts the duration since last strategy change. An agent cannot consider changing strategy again even if the agent is dissatisfied with it during the confirmation period. This is consistent with the theory of innovation diffusion [22] in which an agent has a confirmation period just after adopting a new strategy. Both counters are necessary to simulate an agent's stability and consistency. The corresponding delayed action of both counters can only occur when the counter is above the parameter TD.

Except for W_i, all these attributes of a farmer agent are dynamic during the simulation and are described in detail below.

Credibility Each agent *f* gives a credibility $C(f, v)$ to another agent *v* by comparing their outcomes. Credibility is between 0 (not credible at all) and 1 (very credible).

Satisfaction Each agent has a satisfaction with its current farming strategy (I_C) that corresponds to an evaluation of strategy. It may also evaluate an alternative strategy in certain cases and has a satisfaction for it (I_A). Satisfaction with a farming strategy lies between 0 (not satisfied at all) and 1 (very satisfied).

If an agent is satisfied with its current farming strategy, it does not consider an alternative. Otherwise, its satisfactions with its current farming strategy (I_C) and with an alternative one (I_A) are computed and compared. If I_A is higher enough than I_C, the agent will change its farming strategy. I_C is thus computed at every iteration, whereas I_A is only computed when a stable agent is dissatisfied with its current farming strategy.

In accordance with TRA, the satisfaction I_S with a farming strategy S depends on two elements: attitude A_S and subjective norm SN_S toward S. In the original theory, the interaction between these two elements varies with different agents facing different situations. In order to keep the model simple, satisfaction is assumed as the average value of these two elements.

$$I_S = \frac{A_S + SN_S}{2} \tag{49.1}$$

Both attitude and subjective norm lie between −1 (very negative attitude/subjective norm concerning the farming strategy to evaluate) and 1 (very positive attitude/subjective norm concerning the farming strategy to evaluate). They are computed with farms' outcomes, farmers' strategies, and credibility. See the section "Farmers' dynamics" for the computation details.

Considering the value range of attitude and subjective norm toward a farming strategy, the satisfaction should also lie between -1 and 1. However, to facilitate other calculations, the satisfaction is normalized between 0 and 1.

Practice As stated above, the term "practice" in the model is not really an agent's actual practice, but the way the agent evaluates it (i.e., its performance). A practice is evaluated over two dimensions: the level of output production (i.e., the productivity impact, in our case milk production) and the level of environmental amenity production (i.e., the environmental impact: mineral impacts, energy consumption, among others), respectively, called productivity performance (P_0) and environmental performance (P_1) in the chapter. Both P_0 and P_1 lie between 0 (very bad on this practice dimension) and 1 (very bad on this practice dimension).

Importance given to each dimension The importance given to productivity dimension is termed W_0, and the one given to environmental dimension is termed W_1. W_0 and W_1 lie between 0 (not important at all) and 1 (most important). They sum to 1.

$$W_0 + W_1 = 1 \tag{49.2}$$

Importance defines an agent's personal values. An agent uses its own lens to judge the information it receives and the other agents it meets. In this model, both W_0 and W_1 are kept constant if an agent does not change its farming strategy.

Farming strategy It is defined by the importance that a farmer gives to each dimension of practice. Two farming strategies are considered: organic and conventional. The organic strategy means lending more importance to environmental dimension and less to productivity dimension, whereas the conventional strategy does the contrary. It is assumed that when a farmer agent changes its strategy, it changes accordingly the importance given to each dimension.

Farm

A farm has three attributes: its farming total production (productivity outcome) T_0, its environmental amenities outcome T_1, and its reference R. R is the maximum possible productivity performance considering a farm's all characteristics and evolution. Interviews and experts' arguments show that conventional farms' references are grounded on the negotiations with dairy enterprises (often expressed by "quota" in Europe in the past). Organic farms have more constraints in terms of reference due to stricter regulations. Thus, a conventional farm f's reference is considered as its farmer f's initial productivity performance $P_0{}^f$ $(t = 0)$ and that for an organic farm is a function l of $P_0{}^f$ $(t = 0)$.

$$R^f = \begin{cases} P_0^f\ (\mathrm{t}=0)\ \text{ if } f \text{ is a conventional farm/farmer} \\ l\left(P_0^f\ (\mathrm{t}=0)\right)\ \text{ otherwise} \end{cases} \tag{49.3}$$

The implementation of a farm may need more attributes for different use cases. The detailed computation of T_0 and T_1 is defined in the model's implementation (see 49.2.2).

Media

When an agent is dissatisfied with its current farming strategy and looks for an alternative, it first searches in the population for other agents having similar characteristics but applying an alternative strategy. If it cannot find one, it has access to the media for a stereotype alternative model which depends on the farm's current outcomes.

49.2.1.2 Dynamics

Overview of a Farmer's Dynamic Over the Years

One time step (iteration) t –> $t + 1$ represents 1 year, that is, farmers decide their farming strategies, their practices, and so on, once a year. During an iteration, farmers' update order is picked up at random by a uniform law.

```
For each iteration {
    Generate the order of the population
    For each agent f in the population {
        Compute I_C
        If DC>TD and I_C<TA, compute I_A
            If DD>TD and I_A > I_C + TO, change strategy and update W_0, W_1
        For each agent v that is different from agent f in the population,
        compute C (f, v)
        Compute P_0, P_1 } }
```

Algorithm 1— Population updating loop. I_C is the satisfaction with a current strategy; I_A is the satisfaction with an alternative one. DC is an agent's confirmation duration; DD is an agent's dissatisfaction duration. TD is the minimum time of dissatisfaction before considering the alternative. TA is threshold of I_C to consider an alternative. TO is the threshold of I_A to change strategy. W_0 is the importance given to productivity performance, and W_1 is the importance given to environmental performance. $C(f, v)$ is the credibility that agent f gives to agent v. P_0 is the productivity performance, and P_1 is the environmental performance.

As shown in Fig. 49.1 and the pseudocode of Algorithm 1, during each iteration, an agent evaluates its satisfaction with its current farming strategy. If the agent is

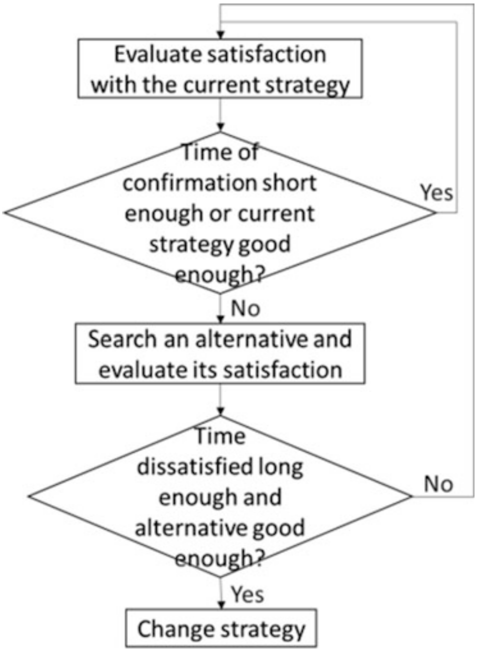


Fig. 49.1 Overview of the farmer's update

in a stable period and is satisfied with its current strategy, it does not consider a change. Otherwise, the agent looks for an alternative and evaluates it. If the agent has been dissatisfied for long enough and the alternative is good enough, it will change. Otherwise, the agent stays with its current farming strategy. It will then update its credibility given to other agents and its practice. See the detail in the following.

Credibility Update

Every two agents' relationship is characterized by the credibility one gives to another and depends on an agent's personal view of its difference to another in outcome (i.e., total production). For agent *f*, its difference to agent *v* is the sum of difference on each outcome dimension weighted by the importance given to that dimension.

$$D_v^f = \sum_{i=1}^{2} \left(W_i^f \left(T_i^v - T_i^f \right) \right) \tag{49.4}$$

The credibility that agent *f* gives to agent *v* is calculated with *f*'s difference to *v*:

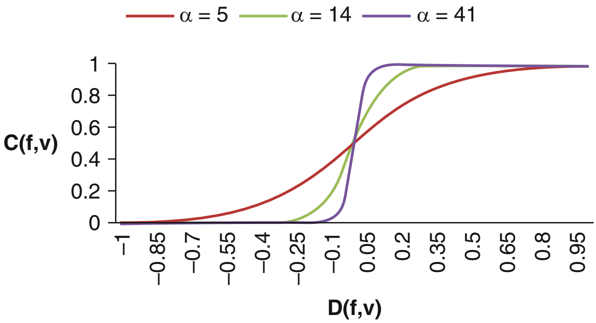


Fig. 49.2 The credibility (*y* axis) agent *f* gives to *v* depends on *f*'s difference (*x* axis) to *v* for three values of parameter α (different-colored lines)

$$C_v^f = \frac{1}{1 + e^{-\alpha \, D_v^f}} \tag{49.5}$$

with α the parameter to characterize the slope of the logistic function.

In Fig. 49.2, agent *f*'s difference and credibility to *v* are, respectively, plotted on the *x* axis and *y* axis. When the difference is negative, it means that *v* has a worse outcome than *f*; thus, *f* gives little credibility to *v*. When the difference is positive, *v* has a better outcome than *f*; thus, *f* gives big credibility to *v*.

The lines with different colors represent α variations to characterize the bias degree that an agent has for others with better outcomes. When α is small, the bias is small. An agent tends to give the same credibility to others, whether or not they reach better outcomes. If α is big, the bias is strong. Only others with better outcomes are credible.

In the model, every two agents are connected. The credibility depends on an agent's perceived difference in outcome to another, and it is then used to update the agent's outcome which can change the perceived difference. Thus, these elements are dynamic.

Farming Strategy Change

An agent changes its farming strategy according to its satisfaction evaluations with its current strategy (I_C) and with the alternative one (I_A). If an agent is in a stable state (its confirmation duration DC > threshold TD) and it is still dissatisfied with its current strategy (I_C < threshold TA), it will consider an alternative one. If the agent is dissatisfied long enough (its dissatisfaction duration since being stable DD > TD) and its satisfaction evaluation of the alternative is better enough than that with its current one ($I_A > I_C$ (1+ threshold TO)), it will change strategy. As stated in Eq. (49.1), satisfaction *I* with a strategy is the average sum of the related attitude *A* and subjective norm SN.

In Eq. (49.1), attitude (A_S) represents an agent's personal view of the difference between its experience and the (potential) outcome of evaluated strategy S. The agent f's experience is its average outcome on the farm ($\overline{T_{C,0}}$, $\overline{T_{C,1}}$) with its current farming strategy (S^f) in memory (M). It is computed as follows:

$$\overline{T_{C,i}} = \frac{\sum_{t \text{ and } S^t = S^f}^{M} T_i^t}{\text{Nb}\left(S^t = S^f\right)} \tag{49.6}$$

The evaluated outcome depends on the strategy to be evaluated. For agent f's current farming strategy evaluation, the evaluated outcome is f's current outcome (T_i^f).

Agent f's attitude toward the current farming strategy (A_C) is as follows:

$$A_C^f = \sum_{i=1}^{2} \left(W_i^f \left(T_i^f - \overline{T_{C,i}^f} \right) \right) \tag{49.7}$$

If an agent's outcome on its farm changes and this change is considered better than its experience, A_C will be positive and strengthen the agent's decision to keep its current strategy. Otherwise, A_C will be negative and may influence I_C. Then, the agent may be dissatisfied and evaluate I_A.

For agent f's evaluation of an alternative farming strategy, the evaluated outcome is the average outcome ($\overline{T_{A,0}^f}$, $\overline{T_{A,1}^f}$) of other agents having similar characteristics as f but applying the alternative. The similarity is defined by a function Y over the farm's characteristics and is compared with a distance threshold (simi). Y is designed in the model's implementation (see 49.2.2). The evaluated outcome is computed as follows:

$$\overline{T_{A,i}^f} = \frac{\sum_{Y_v^f < \text{simi and } S^v \neq S^f}^{N} T_i^v}{\text{Nb}\left(Y_v^f < \text{simi and } S^v \neq S^f\right)} \tag{49.8}$$

If there is no corresponding peer (no other agents similar to f and applying the alternative strategy), agent f will search the media for a stereotypical farm as the alternative.

$$\overline{T_{A,i}^f} = T_{\text{model},i}^f \tag{49.9}$$

Therefore, agent f's attitude toward an alternative is:

$$A_A^f = \begin{cases} 0 \text{ if } \left(I_C^f > \text{TA}\right) \\ \sum_{i=1}^{2} \left(W_i^f \left(\overline{T_{A,i}^f} - \overline{T_{C,i}^f} \right) \right) \text{ otherwise} \end{cases} \tag{49.10}$$

Another component of satisfaction, the subjective norm, represents how an agent considers others' opinions on the evaluated farming strategy through outcomes (i.e., the strategy's implementation results). It is thus an agent's perceived difference between the outcome to be evaluated and the average of other agents' outcomes.

For agent *f*'s evaluation of current farming strategy, the subjective norm is:

$$\mathrm{SN}_C^f = \sum_{i=1}^{2} \left(W_i^f \left(T_i^f - \frac{\sum_{v \neq f}^{N} \left(C_v^f T_i^v \right)}{\sum_{v \neq f}^{N} C_v^f} \right) \right) \tag{49.11}$$

An agent will be socially satisfied if it perceives that other agents, especially those to whom it lends great credibility ("important others"), consider it as a "good farmer." The agent may be so satisfied to have a good social image that it will never consider a major change. Otherwise, if the agent feels socially bad, it may try to become more similar to others in the group or to change of group. This can be done with a change of strategy.

For the evaluation of an alternative farming strategy, the subjective norm is:

$$\mathrm{SN}_A^f = \begin{cases} 0 \; if \; \left(I_C^f > TA \right) \\ \sum_{i=1}^{2} \left(W_i^f \left(\overline{T_{A,i}^f} - \frac{\sum_{v \neq f}^{N} \left(C_v^f T_i^v \right)}{\sum_{v \neq f}^{N} C_v^f} \right) \right) \text{otherwise} \end{cases} \tag{49.12}$$

If in other agents' opinions, especially those to whom agent *f* lends great credibility ("important others"), the alternative is not better, then it is judged not good enough to improve the situation. Agent *f* will tend to keep its current strategy. Otherwise, the agent's subjective norm strengthens its intention to change its strategy.

If an agent changes its strategy, it also changes the importance given to each dimension of practice. The parameter W is the conventional farmers' initial importance given to the productivity dimension.

$$\text{For conventional agents}: W_0 = W;\ W_1 = 1 - W \tag{49.13}$$

$$\text{For organic agents}: W_0 = 1 - W;\ W_1 = W \tag{49.14}$$

Practice Update

As farmers co-construct their practices [20], at each time *t*, a farmer agent updates its practice by copying the performance of its credible peers with a similar farm.

$$\Delta P_i^f = \frac{\sum_{v \neq f \text{ and } Y_v^f < \text{simi}}^{N} C_v^f \left(P_i^v - P_i^f \right)}{\sum_{v \neq f \text{ and } Y_v^f < \text{simi}}^{N} C_v^f} \tag{49.15}$$

Both dimensions of practice are between 0 and 1. A farmer's productivity performance is also limited by the reference R on its farm.

$$P_0^{f,t+1} = \begin{cases} 0\ if\ \left(P_0^{f,t} + \Delta P_0^f < 0\right) \\ R^f\ if\left(\left(P_0^{f,t} + \Delta P_0^f\right) > R^f\ \right) \\ P_0^{f,t} + \Delta P_0^f \text{ otherwise} \end{cases} \tag{49.16}$$

A special case: if an agent f looks for an alternative and cannot find a similar peer applying an alternative strategy, it will look for an alternative in the media. If after evaluation agent f adopts the alternative in the media, it will also copy the practice.

49.2.2 The Design of Use Cases Based on Agricultural Data

The model is implemented on data from agricultural general French census (RGA 2000) about six French "cantons" with strong variations in terms of practice intensity and homogeneity: 1516, Ruynes-en Margeride; 4202, Boën; 4224, Saint-Genest-Malifaux; 2208, Chèze; 2235, Plouagat; 2522, Quingey. The initialization of population distributions in each "canton" is deterministic. For sake of simplicity and according to expertise and literature [23, 24], a farmer's farm and practice are defined by three variables: the utilized agricultural area (UAA), the number of dairy cows (NC), and the quota (Q) which is a synthetic indicator of the farm's maximum milk volume. Figure 49.3 shows the average value of these variables in each "canton."

The degree of "intensity" is measured as the average UAA/average NC. "Cantons" 2208 and 2235 are the biggest producers and most intensive, while 4202,

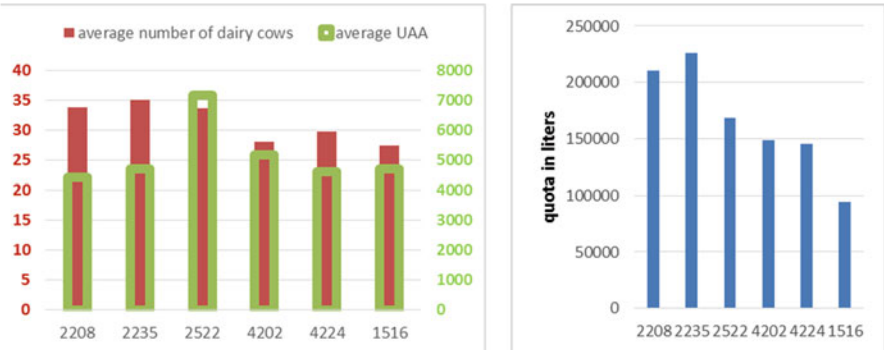


Fig. 49.3 Average number of dairy cows NC, average UAA (in are), and *quota* (in liters) for the six chosen French "cantons"

4224, and 1516 have the least milk production. Despite its average milk production, 2522 is the most extensive "canton" and also has the largest average UAA.

49.2.2.1 Farmer

A farmer agent is designed by its practice with productivity performance P_0 and environmental performance P_1 designed from data. P_0 is directly deduced from the farm's initial characteristics and corresponds to the normalized average milk volume produced by one cow in 1 year. For farmer/farm *f*, at the initial time $t = 0$, $P_0{}^f = Q^f/\mathrm{NC}^f$. At every time *t*, $P_1{}^f = T_1{}^f/\mathrm{SAU}^f$; Eq. (49.15) is only used to update $P_0{}^f$.

The *Y* function telling how two farmers are judged similar is based on a similarity of their farms' characteristics regarding UAA and NC. For agent *f*, agent *v* is a similar peer if $\frac{|\mathrm{UAA}^f - \mathrm{UAA}^v|}{\mathrm{UAA}^f} < \mathrm{simi}$ and $\frac{|\mathrm{NC}^f - \mathrm{NC}^v|}{\mathrm{NC}^f} < \mathrm{simi}$. The threshold simi is a parameter and supposed as 0.1 in the model.

49.2.2.2 Farm

Each farm is initialized by the crossed distribution of discretized utilized agricultural areas (UAA) and quotas (*Q*) of its "canton" from the RGA 2000. They remain constant all along the simulation. The number of dairy cows (NC) is computed from a law extracted from data (with a regression $r^2 = 0.9563$):

$$\mathrm{NC} = 0.2463\ \mathrm{UAA} + 0.0001106\ Q \tag{49.17}$$

From databases regarding farmers' production and various sources,[1] a law is built to compute the potential maximum milk production of an organic farm starting with the conventional strategy, knowing its initial productivity performance P_0 after the normalization and P_0*'* before the normalization. A farm *f*'s normalized reference *R* and reference *R'* before the normalization are computed as follows:

$$\text{For a conventional farm}: R \ = P_0\,(t = 0) \tag{49.18}$$

$$\text{For an organic farm}: R' = 0.6046\ P_0{}'\ (t = 0) + 1913\ \mathrm{NC} \tag{49.19}$$

The environmental amenity outcome T_1 is computed at every time by an aggregated function of literature [23, 24]. It considers mineral impacts and energy consumption related to the total milk production and the farm's agricultural surface:

[1] http://www.cantal.chambagri.fr/fileadmin/documents/Internet/Autres%20articles/pdf/2014/Bio/ABBL2008-2012.pdf
http://www.agrobio-bretagne.org/

$$\text{For a conventional farm}: T_1 \quad = (53 \text{ UAA} + 2.918 \ T_0)/2 \tag{49.20}$$

$$\text{For an organic farm}: T_1 \quad = (-10 \text{ UAA} + 2.588 \ T_0)/2 \tag{49.21}$$

Using French dairy farms' database in RGA 2000, *R'* is to be normalized between 0 (very low production) and 1 (very high production). The 53 UAA, 2.918 T_0, −10 UAA, and 2.588 T_0 are normalized values between 0 and 1. The normalization is:

$$x = \left(x' - \text{min}\right) / (\text{max} - \text{min}) \tag{49.22}$$

where min is the minimum real value in the database, max is the maximum real value in the database, and *x* is the normalized value of real value *x'*.

49.2.2.3 The Media

We use laws extracted from data to design farmers' alternative models. When a conventional farmer *f* wants to evaluate the organic strategy at time $t + 1$, it computes $T_0'(t + 1)$ as follows and its related $T_1'(t + 1)$ with the Eq. (49.20). Note that T_0' is the farmer's real productivity outcome on the farm (before the normalization).

$$T_0'(t + 1) = 0.6046 \ T_0'(t) \quad + 1913 \text{ NC} \tag{49.23}$$

When an organic farmer *f* wants to evaluate the conventional strategy at time $t + 1$, it computes $T_0'(t + 1)$ as follows and its related $T_1'(t + 1)$ with the Eq. (49.21).

$$T_0'(t + 1) = \left(T_0'(t) - 1913 \text{ NC}\right)/0.6046 \tag{49.24}$$

49.2.3 Experimental Design

Reference [14] has studied 625 parameter sets identifying all the qualitative behaviors of the model to diagnose reasons for the absence of conversion. This work aims to study conversion in six different French "cantons" using the same 625 parameter sets varying the main parameters: α (slope of logistic function) from 5 to 41, TA (threshold to consider an alternative) from 0.41 to 0.49, TO (threshold to consider an alternative) from 0.01 to 0.09, and *W* (importance given to the dimension representing farming strategy) from 0.6 to 1. TD (threshold for two counters of duration) is kept constant with the value 5. Agents' memory is also kept constant as 10 years. The distance for similarity, simi, values 0.1. The evolution of each "canton" with a population of 100 conventional farmers (no organic farmers at the initialization) is simulated for 30 years and is replicated 100 times.

49.3 Model Behavior

49.3.1 The Adoption Rate and "Cantons"

At the initialization, each "canton" has no organic farm. Figure 49.4 shows that farms in different "cantons" have different adoption percentages after 30 years. "Canton" 2522 converts the most, followed by 4202 and 1516, and then 4224. 2208 and 2235 convert the least. As these 6 "cantons" have different farm characteristics. The population distribution is studied to learn more about the conversion rate.

49.3.2 The Characteristics of "Cantons"

As shown in Sect. 49.2, each farm has a couple of outcomes (total production) on two dimensions: productivity outcome (T_0) and environmental outcome (T_1). These two values indicate directly a farm's utilized agricultural area (UAA) and quota (Q) which are constant characteristics along the simulation. With another constant characteristic on the farm (i.e., number of cows (NC)), a farmer's initial practice performances (i.e., productivity performance (P_0) and environmental performance (P_1)) are also defined. In Fig. 49.5, for each canton, its average initial T_0 and T_1 are shown on the left; its average initial P_0 and P_1 are shown on the right. The black lines in each histogram show the standard deviation.

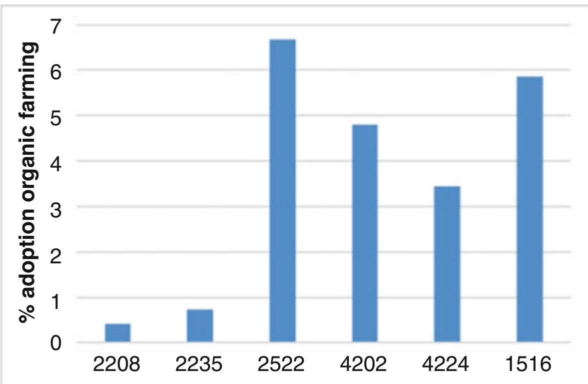


Fig. 49.4 Average percentage of organic farming adoption according to "canton" over 625 sets of parameters and their related 100 replications

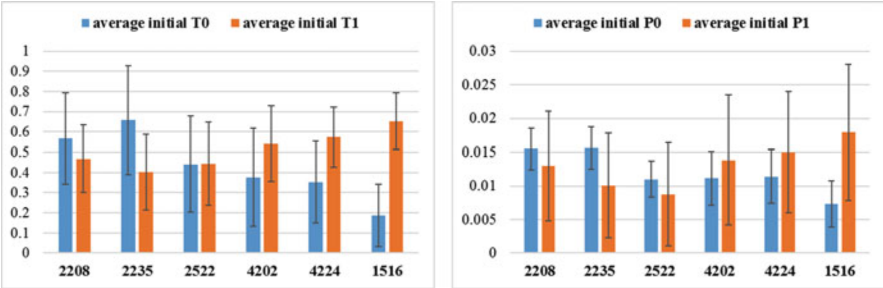


Fig. 49.5 Average of initial T_0 and T_1 (left panel) and P_0 and P_1 (right panel) in each “canton.” Error bars indicates “more or less” one standard deviation

Table 49.1 Initial farm distributions according to “cantons” and farm types in terms of T_0 or T_1

initial T0 \ canton	2208	2235	2522	4202	4224	1516
0.05	3	3	5	16	17	37
0.15	3	2	12	6	5	27
0.25	4	7	14	22	21	17
0.35	8	2	16	14	18	9
0.45	25	21	15	18	12	5
0.55	18	12	16	9	14	2
0.65	11	6	10	3	10	2
0.75	9	3	2	2	2	1
0.85	8	10	2	6	1	0
0.95	11	34	8	4	0	0

initial T1 \ canton	2208	2235	2522	4202	4224	1516
0.85	3	3	3	13	14	19
0.75	5	5	7	5	7	29
0.65	14	5	15	22	21	21
0.55	25	22	16	19	26	14
0.45	24	19	23	24	23	10
0.35	15	7	10	4	6	5
0.25	5	23	11	6	1	2
0.15	7	14	5	4	2	0
0.05	2	2	10	3	0	0

Consistently with Figs. 49.3, 49.4, and 49.5 show that “cantons” adopting the least (i.e., 2208 and 2235) are the most intensive and productive (having the smallest UAA/NC and the largest initial T_0). They are followed by 4224, 4202, and 1516, which have the largest initial environmental outcome T_1 and lower productivity outcome T_0. The “canton” with the most adoption (i.e., 2522) is the most extensive (Fig. 49.3), but has the worst initial environmental performance P_1 (Fig. 49.5). Its outcomes: T_0 and T_1 are medium.

The credibility a farmer gives to another depends on their relative outcomes (T_0 and T_1). The social evaluation of satisfactions depends also on their relative T_0 and T_1 and on the outcomes of its “similar” farmers (Eqs. 49.11 and 49.12). Table 49.1 shows the initial farm distributions of each “canton” in terms of T_0 and T_1.

“Cantons” with the most converted agents (i.e., 2522) have medium average T_0 and T_1, while its distribution is diverse. The other “cantons” have more concentrated distributions. “Cantons” 2208 and 2235 have few farms with low T_0. “Cantons” 4224, 4202, and 1516 have numerous farms with high T_1 and very few farms with low ones.

Table 49.2 Average percentage of organic farming adoption according to "canton" and farm types in terms of T_0 or T_1 over 625 sets of parameters and their related 100 replications

initial T0 \ canton	2208	2235	2522	4202	4224	1516
0.05	0.29	0.36	0.93	2.37	2.26	4.19
0.15	0.06	0.07	1.91	0.68	1.05	1.98
0.25	0.02	0.26	1.91	1.75	0.65	1.28
0.35	0.02	0.01	2.02	0.78	0.32	0.55
0.45	0.06	0.17	1.04	0.72	0.17	0.25
0.55	0.11	0.07	1.4	0.33	0.16	0.15
0.65	0.02	0.01	0.59	0.06	0.12	0.16
0.75	0	0	0.1	0.05	0	0.02
0.85	0	0.01	0.06	0.1	0.03	0
0.95	0.01	0.04	0.23	0.01	0	0

initial T1 \ canton	2208	2235	2522	4202	4224	1516
0.85	0.29	0.36	0.25	1.41	1.39	1.16
0.75	0.08	0.15	0.93	1.07	1.07	2.69
0.65	0.08	0.02	2.12	1.34	0.46	2.27
0.55	0.02	0.27	2.07	1.63	1.48	1.18
0.45	0.01	0.17	1.87	0.98	0.26	0.96
0.35	0.11	0.01	1	0.15	0.07	0.28
0.25	0	0.02	1.31	0.23	0.03	0.05
0.15	0.01	0.01	0.37	0.05	0	0
0.05	0	0.01	0.29	0	0	0

49.3.3 Adoption Rate and Farm Types

The adoption rate varies with "cantons." In fact, it varies also with initial outcome types in terms of productivity outcome (T_0) and environmental outcome (T_1).

Generally, the probability to convert is high for cases with numerous farms at the initialization. Farms with a low initial T_0 or a high initial T_1 (first lines in Table 49.2) convert more than others. However, for the highest initial T_1, the conversion rate is not always the largest (except in 2208 and 2235).

The distributions of adoption vary with "cantons." Canton 2522 has a disperse distribution, and its adoptions center on medium outcomes. For 4202, also having a disperse distribution, its adoptions are rather in cases with high T_1 and low T_0. Cantons 1516 and 4224 have a centered distribution, and their adoptions are rather in cases with very low T_0 and high T_1.

"Cantons" converting the least exhibit distribution profiles with one higher density for medium T_0 and another one for the highest T_0. They are rather homogeneous in T_1.

49.4 Discussion and Conclusion

We argued that the organic farming adoption is strongly influenced by the imitation of "important others" practices and is very sensitive to the distance between farms [14, 15]. However, the model proposed and studied considers only prototypical populations and points out the reasons for non-adoption of organic farming. In this chapter, we focus on the impact of population characteristics on the adoption degree. In this work, the model is explored further by an initialization with six French "cantons" (about 100 dairy farmers) varying in their farm type distributions from the French RGA 2000.

At farm level, we have observed that farms with low total milk production and high total environmental production convert more than others. Such farms are said to have extensive practices and are more susceptible to convert according to experts. Approximately one-third of the global land under organic production is located in unfavorable areas where smallholder farmers, who often lack access to insurance or inexpensive credit, dominate agricultural production [25]. Keeping their extensive practices, or even increasing this degree of intensity, by adopting, for example, organic farming, can be beneficial to eliminate their reliance on expensive, fossil fuel-derived chemical inputs [26, 27]. This increasing autonomy is taken into account in our environmental outcome which considers the decreasing of inputs and energy consumption.

At the regional level, intensive regions convert less than extensive ones, especially those with not very good environmental outcomes. Indeed, we notice that the most intensive "cantons" almost do not convert (those located in the Brittany region). They are more likely to be satisfied with their current situation and do not consider a major change. Despite of having the least productivity outcomes, "cantons" 1516, 4202, and 4224 do not have the most conversion because of their already good environmental outcome. The adoption of organic farming cannot bring them much environmental gain. On the contrary, the most extensive "canton" (i.e., 2522) has the largest adoption rate. It is less productive than farms in the Brittany region (2238 and 2235), but more than other "cantons." Its environmental production is not as good as the other extensive "cantons" (1516, 4202, and 4224). Thus, most farms from this "canton" can poorly decrease their productivity outcome and increase much their environmental outcome with the conversion to organic farming. This result can sound strange, but it has been pointed out by Sutherland that most farmers having already chosen low-input strategy (i.e., strongly extensive) do not convert since they cannot improve more their environmental impact [28]. The only reason they can convert is for economic purpose in case of strong crisis. We also show that more diverse regions, having high densities for different farm types, tend to adopt more, but this point deserves more investigations. In fact, our model can also be applied to other contexts, for example, other forms of adoption in agriculture.

However, we cannot compare our adoption rates for the chosen French "cantons" for the following reasons: firstly, because the data are not so easy to obtain and secondly, very important dynamics are not present in the model (the incertitude, difficulty, etc.) and make the comparison having no sense. Adding to the model external factors such as economic crisis, but also demographic evolution implying an increasing average farm sizes, are the next steps for research about this model.

Acknowledgments We acknowledge funding from the Scientific Research Foundation of Jiangxi University of Science and Technology (Grant No.JXXJBS19012).

References

1. Dedieu, M.-S., Lorge A., Louveau O., Marcus V.: Les exploitations en agriculture biologique: quelles performances économiques? (2017)
2. Sainte-Beuve, J., Bougherara, D., Latruffe, L.: Performance économique des exploitations biologiques et conventionnelles: Levier économique à la conversion, 10p Transversalités L'agriculture Biol. Strasbg. (2011)
3. Sutherland, L.-A., Burton, R.J.F., Ingram, J., Blackstock, K., Slee, B., Gotts, N.: Triggering change: towards a conceptualisation of major change processes in farm decision-making. J. Environ. Manage. **104**, 142–151 (2012)
4. Rickards, L., Howden, S.M.: Transformational adaptation: agriculture and climate change. Crop Pasture Sci. **63**(3), 240 (2012)
5. Dowd, A.-M., Marshall, N., Fleming, A., Jakku, E., Gaillard, E., Howden, M.: The role of networks in transforming Australian agriculture. Nat. Clim. Chang. **4**(7), 558–563 (2014)
6. Fairweather, J.R., Hunt, L.M., Rosin, C.J., Campbell, H.R.: Are conventional farmers conventional? Analysis of the environmental orientations of conventional New Zealand farmers. Rural. Sociol. **74**(3), 430–454 (2009)
7. Darnhofer, I., Schneeberger, W., Freyer, B.: Converting or not converting to organic farming in Austria:farmer types and their rationale. Agric. Hum. Values. **22**(1), 39–52 (2005)
8. Rigby, D., Young, T., Burton, M.: The development of and prospects for organic farming in the UK. Food Policy. **26**, 599–613 (2001)
9. Pavie, J., Dockès, A.-C., Echevarria, L.: Etude des freins à la conversion à l'agriculture biologique des exploitations laitières bovines. Paris Cedex 12: Institut de l'Elevage, (2002)
10. Barbier, C., Cerf, M., Lusson, J.-M.: Cours de vie d'agriculteurs allant vers l'économie en intrants: les plaisirs associés aux changements de pratiques. Activités. **12**(12–2), (2015)
11. Burton, M., Rigby, D., Young, T.: Modelling the adoption of organic horticultural technology in the UK using duration analysis. Aust. J. Agric. Resour. Econ. **47**(1), 29–54 (2003)
12. Goldstone, R.L., Janssen, M.A.: Computational models of collective behavior. Trends Cogn. Sci. **9**(9), 424–430 (2005)
13. Grimm, V.: Ten years of individual-based modelling in ecology: what have we learned and what could we learn in the future? Ecol. Model. **115**(2), 129–148 (1999)
14. Xu, Q., Huet, S., Poix, C., Boisdon, I., & Deffuant, G. (2018). Why do farmers not convert to organic farming? *Modeling conversion to organic farming as a major change. Natural Resource Modeling., 31*, e12171. https://doi.org/10.1111/nrm.12171
15. Huet, S., Rigolot, C., Xu, Q., De Cacqueray-Valmenier, Y., Boisdon, I.: Toward modelling of transformational change processes in farm decision-making. Agric. Sci. **09**(03), 340–350 (2018)
16. Kaufmann, P., Stagl, S., Franks, D.W.: Simulating the diffusion of organic farming practices in two New EU Member States. Ecol. Econ. **68**(10), 2580–2593 (2009)
17. Ajzen, I., Fishbein, M.: Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research. Addison-Wesley, Reading (1975)
18. Mintzberg, H., Raisinghani, D., Theoret, A.: The structure of 'unstructured' decision processes. Adm. Sci. Q. **21**(2), 246 (1976)
19. Öhlmér, B., Olson, K., Brehmer, B.: Understanding farmers' decision making processes and improving managerial assistance. Agric. Econ. **18**(3), 273–290 (1998)
20. Darré, J.-P.: La parole et la technique: l'univers de pensée des éleveurs du Ternois, vol. 10. Editions L'Harmattan, (1985)
21. Grimm, V., et al.: ODD protocol : a standard protocol for describing individual-based and agent-based models. Ecol. Model. **198**(1–2), 115–126 (2006)
22. Rogers, E.M.: *Diffusion of Innovations*, 3rd edn. Free Press; Collier Macmillan, New York/London (1983)

23. Pavie, J., Chambaut, H., Moussel, E., Leroyer, J., Simonin, V.: Evaluations et comparaisons des performances environnementales, économiques et sociales des systèmes bovins biologiques et conventionnels dans le cadre du projet CedABio. Renc Rech Rumin. **19**, 37–40 (2012)
24. Chambaut, H. et al.: Profils environnementaux des exploitations d'élevage bovins lait et viande en agriculture biologique et conventionnelle: enseignements du projet CedABio, Rencontres Autour Rech. Sur Rumin., pp. 53–56, 2011
25. F. IFOAM, "The World of Organic Agriculture: Statistics & Emerging Trends 2017." (2017). https://www.ifoam.bio/en/news/2017/02/09/world-organic-agriculture-2017
26. Pimentel, D., Hepperly, P., Hanson, J., Douds, D., Seidel, R.: Environmental, energetic, and economic comparisons of organic and conventional farming systems. Bioscience. **55**(7), 573 (2005)
27. Uematsu, H., Mishra, A.K.: Organic farmers or conventional farmers: where's the money? Ecol. Econ. **78**, 55–62 (2012)
28. Sutherland, L.-A.: 'Effectively organic': environmental gains on conventional farms through the market? Land Use Policy. **28**(4), 815–824 (2011)

Index

H. Verhagen et al. (eds.), *Advances in Social Simulation*, Springer Proceedings in Complexity, https://doi.org/10.1007/978-3-030-34127-5

B

C

L

M

N

T