Corinna Elsenbroich · Nigel Gilbert

Modelling Norms

Springer

Modelling Norms

Corinna Elsenbroich • Nigel Gilbert

Modelling Norms

Springer

Corinna Elsenbroich
Department of Sociology
Centre for Research in Social Simulation (CRESS)
University of Surrey
Guildford, UK

Nigel Gilbert
Department of Sociology
Centre for Research in Social Simulation (CRESS)
University of Surrey
Guildford, UK

ISBN 978-94-007-7051-5 ISBN 978-94-007-7052-2 (eBook)
DOI 10.1007/978-94-007-7052-2
Springer Dordrecht Heidelberg New York London

Library of Congress Control Number: 2013943121

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Acknowledgements

This book took a little longer than expected but books, it seems, have their own schedule, independent of that of the authors, publishers or funders. It is impossible to thank everyone who has contributed directly or indirectly to the work but some people should be explicitly mentioned. The writing of the book was funded by the UK Economic and Social Research Council through the National Centre for Research Methods within a project called Simulation Innovation: A Node. One of the aims of the project was to develop and disseminate agent-based modelling as a method in the social sciences, and that is also the aim of this book, specifically in relation to modelling norms. We thank the ESRC for funding this exciting and important project. As in all projects, many things in addition to research need to be done and we thank Lu Yang especially for handling the organisational issues for us so competently.

We also thank the editorial team at Springer for their patience and help in copy editing the book.

Some people have contributed rather directly to the book through collaboration. We thank Maria Xenitidou for her expertise in social psychology, Piter Dykstra for the development of the model in Chap. 12 and Harko Verhagen for collaboration about collective intentionality. We thank Rainer Hegselmann for commenting on a draft. We are grateful to our colleagues in CRESS, the Centre for Research in Social Simulation, for their companionship and encouragement. Our thanks also go to two anonymous referees whose comments improved the book a lot. All remaining faults and inaccuracies are our own.

Contents

Chapter 1
Introduction

Man is a mediocre egoist; even the most cunning thinks his habits more important than his advantage.

Friedrich Nietzsche

"Life is what you make it." Although we probably often feel creatures of circumstances, this proverb contains an important insight into the human condition: humans have agency, humans can make decisions about their actions. But how do humans make decisions? What motivates them? Why do different people make different decisions in very similar circumstances? And why do people make similar decisions in vastly different circumstances? These questions about human motivations, decisions and behaviours, and the role of circumstances, lie at the heart of the social and behavioural sciences.

The research question "Why do some people commit crimes?" asks why some people break the rules of their own society. Asking this question seems to suggest that non-criminal behaviour is a kind of default that needs no explanation. The question "Why do people cooperate?" asks why people behave in a way that might be detrimental for themselves but beneficial for society. The question seems just as legitimate as the question about crime but suggests that defection, or criminal behaviour, is the default. Which is the right question to ask? Which behaviour is the default? The answer is 'neither'. Both questions are perfectly legitimate. The interesting thing about human behaviour is that *all* behaviour needs to be explained. What is seen as default behaviour and what as explanandum depends solely on the vantage point of the research(er).

Both vantage points above relate to the topic of social norms or normative behaviour. Why people commit crimes, i.e. contravene social norms, is a question of criminology. Why people adhere to norms is a question of sociology/social psychology. In the end these two questions can be combined into one: the question of human behaviour in relation to *social norms*, i.e. either adhering to them or contravening them. This overarching question lies at the heart of the social sciences and is called *inter alia* the structure-agency problem, the micro-macro problem or

C. Elsenbroich and N. Gilbert, *Modelling Norms*, DOI 10.1007/978-94-007-7052-2_1,

the individual-collective behaviour problem. We look at the question of social norms from both sides, discussing theories and models of cooperation in Chap. 2 and crime in Chap. 3.

But first of all we introduce social norms, run through some methodological issues around studying micro and macro phenomena in the social sciences and discuss the possibility of analytical social science.

1.1 Social Norms

Social norms govern most of our life. Although we might be conscious of some norms, like queuing politely for the bus in England, most of our behaviour is relatively automatic, like getting dressed before leaving the house. We would not think of transgressing certain norms, whilst we delight in transgressing others. We would not spit into a stranger's face unprovoked or wear our underpants on our head. But we might delight in crossing a red pedestrian light in Germany where it is the norm to wait at the lights, even if there is no car in sight. So, why do we do what we do? And what are the consequences of many people doing as they do?

This is what this book is about. We will look at questions of deliberate action, social influence, conformity, obedience and compliance and how individual behaviours affect social outcomes and vice versa. We will discuss these questions first from a theoretical perspective in Chaps. 2 and 3, introduce the method of agent-based modelling in Chap. 4 and spend the remainder of the book looking at models explaining different aspects of normative behaviour.

Why do people do what they do? The first answer to this question is usually, 'because they want to'. Although this might sound a little flippant, this answer is a serious contender as an explanation. It is the explanation favoured by individualism, although it is usually phrased in rather more formal terms. Individualist approaches to the question of human behaviour start from two premises: (a) individuals have preferences and (b) individuals try to maximise their own utility. Preferences are simply preferences for certain states of the world over other states. I prefer to have cake to being hungry. I also prefer cake to an apple although I prefer the apple to being hungry. Given these preferences I will act in such a way as to get the cake. If cake is impossible I will try to get the apple. Individualism asserts that agents will do whatever guarantees the best outcome for themselves, they are selfish utility-maximisers. Intuitively this makes sense and it can explain a lot of human behaviour, particularly the bad bits. However, some behaviour is rather more tricky to explain within this framework. Let's try out some examples:

Why do people pay on public transport? On many public transport systems one can get away without buying a ticket. Ticket inspections are relatively rare so that paying the fine if one is caught usually amounts to less than paying every time one takes a ride. Buying a ticket is detrimental to an agent's utility and yet, most people buy tickets.

Why do people give money to charity? Charitable organisations receive a lot of donations from all strata of society. Whilst we might say that rich people can easily give some money away, poor people also donate despite it costing them a lot more in relative terms. There is usually no public recognition for the donation nor any direct benefit to the donator. Donating to charity is detrimental to an agent's utility and yet, many people donate.

Why do people join a trade union? Joining a trade union means paying a membership fee. Trade unions bring advantages such as negotiated pay and work conditions. But workers that are not members of a trade union also benefit from the negotiated pay and work conditions. So, although trade unions provide a benefit for their members, non-members benefit from the trade union negotiations and do not have to pay a membership fee. Thus joining a union is detrimental to an agent's utility and yet, many people join unions.

People seem to do things although they are detrimental to their personal utility. Different explanations are put forward as to why they do. People might buy a transport ticket as it is embarrassing to be caught without. People might donate to charity as it makes them 'feel good'. People might join the trade union because they feel the duty to contribute rather than free-ride. These sentiments are not explicable by individualism.

So people might still do what they want to do but what they want to do seems to be more complicated than simple personal preferences and utility maximisation. An alternative explanation of human behaviour is a structuralist approach in which human action is explained by the social structures in which it occurs. The embarrassment we feel being caught, the joy we feel helping others, the duty we feel to contribute, come from the social structures we live in and we have been socialised into. This structuralist approach, although explaining a lot of social behaviour, has a terrible side-effect: Suddenly people no longer do what they want to do but they are made to act by structures. People lose their agency in the mire of social structures that constrain behaviour. This can be called the *Paradox of Agency*, that on the one hand, individual preferences are not enough to explain the range of behaviours found in the world but social structures seem too coercive, leaving no space for any individual agency.

Our second question was, what are the consequences of many people doing as they do? If all your colleagues are members of the union, you might be more inclined also to join the union. If everybody cancelled their membership of the union, the union would cease to exist. The old slogan 'Imagine There Is a War and No One Turns Up!' comes to mind. Many people behaving in the same or similar way brings about social phenomena such as social institutions (trade unions, war), fashions and fads (suits and ties), traditions (Easter egg hunts, birthday parties, Sunday roasts), conventions (driving on the right side of the road) and even new objects (money, wedding rings, crowns). These behavioural regularities are often called 'social norms'.

We will encounter many definitions of a social norm in the following chapters. We will see that all the definitions capture something relevant about social norms but we will also see in Chap. 10 that each definition fails to cover certain other aspects.

One reason for this difficulty in pinning down a definition of social norms is that norms function on different levels. A social norm might define a society (national stereotyping), demarcate 'groups' within a population (Mods and Rockers), be used as an instrument of power and coercion (Patriarchy), be the reason for an individual's action (conformity, compliance), or many other things. It makes a difference in the definition of a social norm whether we apply it to an aggregate, where it is a behavioural regularity, or to an individual, where it is a rule of conduct.

Social norms also need to be contrasted to moral and legal norms. All three provide rules of conduct but the source of the rules is subtly different.

Definition 1.1.1. A social norm is a rule of conduct derived from a social behavioural expectation.

Definition 1.1.2. A moral norm is a rule of conduct derived from a moral value.

Definition 1.1.3. A legal norm is a rule of conduct derived from the code of law.

The source of moral rules is a set of moral values, the source of legal rules a set of laws. Is it fair to contrast these two sources to social expectations? After all, moral rules are not free from social influence and what is the code of law if not enshrining social values? Although there is certainly social influence on laws and mores, the social influence is not direct as in the case of social norms. The immediacy of social influence can be used as a distinguishing mark of normative behaviour (Elsenbroich and Xenitidou 2012).

1.2 How to Study Social Norms

The study of norms has become more and more interdisciplinary over the past century. Whilst it used to be the prerogative of moral philosophy, psychology and sociology, norms are now a prevalent concept in economics, politics, computer science, logic, artificial intelligence, communication theory and evolutionary biology. Neo-Darwinian theories explain altruism as a phenotype of the selfish gene so that altruism becomes a mechanism that has evolved to increase the chances of gene survival through the generations (Dawkins 1976). Psychological research on norms reaches from assessing hormonal influences on behaviour (Fairchild et al. 2008) to motivation and identity studies (Turner 1996). In computer science norms have informed the subfields of deontic logic, argumentation theory, Artificial Intelligence, human-computer interaction, etc. In economics, norms have been used to explain (away) preferences but also to widen the scope of the *homo economicus* (Ellickson 1998).

In this book we are concerned with a particular methodology applied to the study of social norms, the methodology of agent-based modelling. Agent-based modelling has possibly been the most radical methodological revolution in the social sciences of the past 50 years. We will discuss agent-based modelling in more detail in Chap. 4. It is sufficient for now to know that agent-based modelling

is a computer simulation method for the social sciences. It allows the dynamic representation of (social) processes. An agent-based model can be used to explore experimentally the influences and interdependencies of different parameters on a (social) phenomenon.

1.3 Theoretical Social Science

How can computer simulations help us understand anything about the real world? The social sciences, like the natural sciences, should be based on empirical findings. Comte, grandfather of sociology, called the new science 'social physics' and Durkheim saw it as the natural extension of the natural sciences into the realm of the social. Empiricism and data are important in the natural sciences, at least since Bacon's *Novum Organon*. Nowadays, natural scientists rarely go out into the real world to collect data, rather they stay in and generate data under highly controlled conditions in laboratories. This process is called experimentation and it is the backbone of natural science. The social sciences, on the other hand, are dominated by data gathering in the real world. Whether using interviews, focus groups, surveys, observations, or text analysis, social scientists collect data rather than generate it. Nevertheless, although scarce, experimentation does exist in the social sciences. Experimental economics mainly investigates the soundness of assumptions of game theory, asking subjects for their decisions in artificial social setups. Social psychology conducts experiments into the social influences between people.

However, although there is (some) experimentation at the micro level, i.e. small scale, often dyadic, interactions between people, there is no macro experimentation. There are many reasons for this. There are ethical reasons not to experiment on human beings.[1] The experimentation on humans in concentration camps during the second world war led to a code of ethics enshrined in the Nuremberg Code, containing ten points covering minimal mental and physical suffering, informed consent and maximal reduction of risk. A second problem for the social sciences is the reflexivity of the subjects of study. For example informed consent poses a problem for the social sciences as disclosure of the experimental setup might change the behaviour of people. Had Milgram (1973) informed his 'teachers' that no actual electro shock was administered to the subject, could the experiment have shed any light on human behaviour?[2] A third problem is that social systems are too large and/or too complex to experiment on. A society just does not fit into a laboratory. (For a more detailed discussion, see for example Hollis 1994, Chaps. 3 and 9).

Do the social sciences need experimentation? The answer to this question depends on what is seen as the goal of the social sciences. For some the purpose

[1] See for example http://listverse.com/2008/09/07/top-10-unethical-psychological-experiments/

[2] See Sect. 2.2.4 for a summary of the experiment.

of the social sciences is to represent the social world, to gain (deep) understanding of the meanings of social interactions and to provide a 'thick description' of human behaviour, Geertz (1973). For others the aim of social science is generalisation, perhaps not to the point of finding laws, but at least statistical general statements. Corresponding to these are two kinds of methods. Qualitative analysis, focused on deep understanding of specific small scale samples of social actors and phenomena but not focused on explanation or generalisation, and quantitative methods, mainly statistical analysis, focused on explanation and generalisation.

A third conception of the aim of the social sciences is analytic sociology (Hedström and Swedberg 1998). Here the goal of the social sciences is to find the underlying mechanisms of social phenomena. Rather than staying at the level of specific description or statistical generalisation, the aim is to uncover the 'nuts and bolts' of the social world, to extract the causal relationships (Elster 1989).

Experiments are vital for the investigation of mechanisms. Social scientists using quantitative methods often make causal claims such as "socio-economic deprivation *causes* low attainment in pupils". However, they do not explain low attainment, they merely present evidence for a link between the variables "socio-economic deprivation" and "low attainment". Elucidating a mechanism means to open the black box of a statistical correlation and show *how* an effect is produced. We cannot gain causal knowledge from statistical knowledge directly, or, as Cartwright (1989) puts it, "no causes in, no causes out". And whilst statistical analysis gives us associations of variables over a static population, it does not tell us what happens in a dynamic setting, or statistical "[a]ssociations characterize static conditions, while causal analysis deals with changing conditions." (Pearl 2003, p. 104).

If the social sciences are seen as descriptive and interpretative there is no need for experimentation; observations and recording of data is sufficient. If the goal of the social sciences is, however, to investigate the mechanisms underlying social phenomena, experimentation becomes paramount. In this book we will discuss computer simulation as an appropriate method for the investigation of mechanisms. Computer modelling sounds a lot like armchair philosophising rather than the empiricism a proper science demands. In the following section we look at the use of 'armchair science', i.e. not leaving one's armchair but conducting science rather than 'mere' philosophising. The traditional method of armchair science is thought experimentation. As we will see, thought experimentation has been very important to the natural sciences and we argue that computer simulation is a form of thought experimentation for the social sciences.

1.3.1 Thought Experiments

Rather surprisingly, armchair science has often been at the heart of scientific advances. The instrument of armchair science is the thought experiment, a form of 'what-if' reasoning. Kühne (2005) points to the difference between the "old", intuitive method of explaining the world employed by Aristotle (e.g. the Aristotelian

Law of Free Fall), which Kühne describes as 'picturing nature', to the 'new' method which tries to 'reconstruct nature' using idealisation, non-intuitive premises and superimposition of simpler facts (e.g. Galileo's law of inertia). Traditional social science is similar to the Aristotelian method whereas analytic sociology advocates a paradigm shift towards reconstruction.

As the name *thought experiment* suggests, a thought experiment is an experiment happening purely in thought, or in the 'laboratory of the mind' as Brown (1991) catchily calls it. Examples of thought experiments run through the ages of scientific enquiry, from Galileo's tied together stones, Newton's Bucket, Einstein's Train to Schrödinger's Cat.

For example Galileo's thought experiments on motion undermined prevailing paradigms (although they were not accepted until much later). The then common conception of a body's motion originated from Aristotle's natural philosophy which proposed that bodies have *natural speed*, with heavier bodies moving faster than lighter bodies. But what happens if two bodies are tied together? Galileo (1628/2010) executes a thought experiment in a dialogue between Salvati and Simplicio

> Salvati: If we take two bodies whose natural speeds are different, it is clear that on uniting the two, the more rapid one will be partly retarded by the slower, and the slower will be somewhat hastened by the swifter. Do you not agree with me in this opinion?
>
> Simplicio: You are unquestionably right.
>
> Salvati: But if this is true, and if a large stone moves with a speed of, say, eight, while a smaller stone moves with the speed of four, then when they are united, the system will move with a speed of less than eight. Yet the two stones tied together make a stone larger than that which before moved with a speed of eight: hence the heavier body now moves with less speed than the lighter, an effect which is contrary to your supposition. Thus you see how, from the assumption that the heavier body moves faster than the lighter one, I can infer that the heavier body moves more slowly... And so, Simplicio, we must conclude therefore that large and small bodies move with the same speed, provided only that they are of the same specific gravity.

This thought experiment gives us a good idea of how rationalisation alone can convey knowledge about empirical matters. From the hypothesis that two bodies have different speeds, Galileo asks the question what happens if they are tied together. If the two bodies have different *natural speeds* the slower body has to slow down the heavier body. But the tied together bodies constitute a new single body, heavier than either of the original bodies, which, according to the theory of *natural speeds* has to fall faster. This contradiction shows that the theory of *natural speeds* cannot be true.

The history of science is full of breakthroughs resulting from what can be seen as pure armchair science. This is somewhat at odds with the idea of an experimental science in which nature is the arbiter of truth, not reasoning, and thought experimentation is not without its critics. One major debate is between a Platonic (Brown 2004a,b) and an empirical interpretation of thought experiments (Norton 2004a,b). Brown's position is that thought experiments actually reveal nature because of the possibility of the mind to "peek at the platonic heavens" (Brown

2003). Empiricism is not enough to understand nature making rationalisation an essential part of scientific methodology. Norton's position on the other hand is that thought experiments are "nothing but" arguments. This means that thought experiments do not contain new informative content but are purely deductive. This stands in contrast to what Kühne (1995) calls the 'paradox of thought experiments', as it seems to be the case that they do in fact lead to new empirical insights meaning they cannot be deductive.

Brendel (2004) argues against Brown's Platonism and holds with Norton that thought experiments can be *reconstructed* as arguments, but that their function is more than just that of an argument. Instead, in their execution they are what Dennett (1995) calls *intuition pumps*. The difference between a formal argument and an intuition pump is that the latter does not have to state and reveal all the premises necessary to reach a certain conclusion. On the contrary, as Dennett (1984) states "intuition pumps are cunningly designed to focus the reader's attention on 'the important features, and to deflect the reader from 'bogging down in hard-to-follow details'." This omission of 'irrelevant premises' is on the one hand the strength of thought experiments, as they tap into intuition directly, but at the same time a danger as this omission can lead to abuse.

What makes a good thought experiment? Einstein and Infeld (1938) cite Galileo's thought experiment of a cart on a frictionless plane to show uniform motion in the absence of external forces. This thought experiment superimposes the single factors of force and motion and reconstructs phenomena like friction and inertia. Intuitively, a thought experiment can be described as a hypothetical reconstruction of reality by singling out simple features and superimposing them, reading off their interactions.

The discussion on thought-experiments is often divided into thought-experiments in the natural sciences (Brown 1991; Norton 2004b; Kühne 2005) and thought-experiments in philosophy (Cohnitz 2006); the former clearly juxtaposed to the experimental natural sciences, the latter more in line with an interpretation of thought-experiments as (counterfactual) arguments.

1.3.2 Thought Experiments in the Social Sciences

As we have seen above, there is plenty of high quality literature on thought experiments in the natural sciences (Kühne 2009 for an excellent overview and Brown 1991; Kühne 1995, 2005; Norton 2004b; Brendel 2004 for further philosophical discussion). Another field of study in which thought experiments are prevalent is philosophy itself, in particular in ethics (e.g. survival lottery, Harris 1975), philosophy of mind (e.g. brain-in-vat, Putnam 1982), philosophy of language (e.g. the Chinese Room, Searle 1984) and questions of material or personal identity (e.g. the ship of Theseus, Neurath 1921).[3] The social sciences have thought

[3]For a full analysis of thought experiments in philosophy see Cohnitz (2006).

experiments such as Hume's *Specie-Flow* economy, the Prisoner's Dilemma, Rawls' *Veil of Ignorance* and Hardin's *Tragedy of the Commons*, to name some famous ones, but they are few and far between. We will discuss two in more detail as they exemplify the common kinds of thought experimentation in the social sciences.

1.3.2.1 Hume's Specie-Flow Mechanism

Hume (1752) presents a thought experiment about the flow of money (specie) between economies. Imagine four fifth of the money in Britain to be destroyed over night. What would happen to the economy? Hume argues that commodity and labour prices would fall sharply to accommodate the lack of currency. That in turn would mean that other nations would increase buying British goods as they were comparatively cheap and within due course money would flow back into the British economy. This in turn would lead to increased commodity and labour prices until the trading economies are in equilibrium again. Similarly, if money was multiplied fivefold overnight, prices would increase sharply so that goods from other nations would become more and more attractive, leading to the money flowing out of the British economy.

Hume concludes his thought experiment as follows:

> Now, it is evident, that the same causes, which would correct these exorbitant inequalities, were they to happen miraculously, must prevent their happening in the common course of nature, and must for ever, in all neighbouring nations, preserve money nearly proportionable to the art and industry of each nation. *Of the Balance of Trade* (Hume 1752)

The purpose of the scenario is to show that it is not the absolute quantities of money that matter in an economy but the relative proportions of prices and money and to point to the importance of industry and people rather than the importance of money.

1.3.2.2 Hardin's Tragedy of the Commons

The Tragedy of the Commons (Hardin 1968) is a well known application of rational choice theory to a common good. Imagine a common, a land for everyone to graze their livestock on. One day a herdsman thinks that he should add an animal to his stock. After all, more animals will bring him more money. The herdsman is aware of the possibility of overgrazing but reasons that the increase of that risk by him adding one animal is minute, a fraction of the utility gained by the added animal. The problem is that each and every herdsman should reason this way, thus rapidly leading to overgrazing.

This scenario explores the interdependence of individual utility and common good. The scenario has rational choice theory as a background, asserting that each herdsman would make the decision to add one animal after another to his

herd. Without this assumption the tragedy of the commons vanishes. Hardin uses the tragedy to argue against applications of commons in any area of life and for regulation and private ownership.

Hume's specie-flow describes the interdependence of money, prices and wages. It stays at the level of institutional facts and their interaction without having to touch on human decision making. This kind of thought experiment is mainly used in economics where institutions are more clearly defined than in other social sciences.

Hardin's Tragedy of the Commons explores the interdependence of individual utility maximisation and social utility maximisation. The decision that maximises the individual's utility leads in the long run to a collapse of utility for all. This thought experiment is an example of many relying on rational choice and game theory.

The two kinds of social science thought experiments can be summarised as either analysing the interdependence of institutions or exploring the macro effects of individual decisions. But what happens when we add tariffs to Hume's flow of money? Exchange rates? Cartels? Debt? Soon the analysis of institutional interdependencies becomes too complex to execute in the mind. Similarly for the case of the other kind. In order to get interesting results from rational choice theory or the prisoner's dilemma we need large numbers of agents repeatedly interacting. We can no longer execute experiments in thought, instead we analyse scenarios mathematically, with differential equations and equilibria. These in turn no longer lay open what the causal mechanisms underlying the phenomena are and we are back to square one where we can no longer uncover mechanisms.

1.3.3 Thought Experiments and Agent-Based Modelling

So the main barrier to thought experiments in the social sciences, at both the institutional facts and the individual facts levels, is the complexity of the interactions needed to generate interesting social phenomena. In this section we introduce agent-based modelling as a method of thought experimentation for the social sciences. We have discussed two kinds of thought experiments in the social sciences in Sects. 1.3.2.1 and 1.3.2.2, one investigating the interaction of institutions, the other based on simple scenarios built on rational choice theory. These two examples showed that only relatively simple interactions can be modelled in thought experiments before they quickly become too complicated to think through. Let us see how we can enhance the 'laboratory of the mind' with the virtual laboratory of computer simulations.

Agent-based modelling has been a ground breaking methodological innovation, introducing a new, model-centred epistemology into the social sciences. Agent-based models are computer programs consisting of a number of autonomous, heterogeneous agents interacting with an environment and with each other. The environment in an agent-based model can represent a geography of resources (e.g. houses, food), a social space (e.g. networks), or something more abstract like an

opinion or information space. Agents move about in space, both changing the space (e.g. consuming resources) and reacting to it (e.g. moving towards the nearest resource). Agents also interact with each other, e.g. exchanging information or imitating one another. Given these ingredients, many macro phenomena can be generated from the interactions of agents at the micro level. Agent-based models have successfully been employed in the study of complex social phenomena including opinion dynamics, technology adoption, markets, social networks and segregation.

Agent-based models allow the superimposition of several single facets of reality for reconstruction (Stöckler 2000), making them more similar to thought experiments than other simulations. This somewhat simplifies the justification and validation of agent-based models. Whilst other simulation methods (see Chap. 4) define a complete system in advance and compare the simulation outcomes to empirical data, agent-based-simulation models might be constructed by step-by-step enrichment (Gilbert and Troitzsch 2005). Starting from very few parameters, other factors can be added one by one and at each point, the results can be scrutinised for validity (Gilbert 2008). This reflective process of enrichment allows maximal control over parameters thus preserving intellectual surveyability.

A well-known example of an agent-based model is a model of ethnic segregation discussed in Schelling (1971) (see Chap. 4 for more detail). Schelling wondered about the persistence of segregation in American cities. Imagine agents randomly allocated to the patches on a grid. The agents come in two colours, red and green. Agents have a threshold for how many neighbours (the adjoining patches) of the other colour they tolerate. If the number of other coloured agents exceeds this threshold, the agent moves to an arbitrary other patch on the grid. Schelling initially executed this set of agent interactions on a checkered board by hand. He found that segregation resulted even with agents having very weak preferences for similarity.

There are many computer implementations of this thought experiment as agent-based models (for one example see Wilensky 1997). Although even the initial experiment executed by hand on a checkered board showed the rough results of segregation resulting from weak similarity preferences, in the computer implementation we can vary the parameters (such as the density of agents, the number of neighbours considered and the agents' tolerance thresholds) and see the respective influences of these variables and their interdependencies. For example we can identify a 'tipping point' of the tolerance threshold at about 30 % when the degree of segregation shoots up from the initial 50 % to between 75 and 80 %. Also, segregation gets worse as the density of agents decreases. This is a good example of a virtual experiment. There is a limited number of variables the interaction of which is investigated by systematic variation.

There will never be a city where the decision to move is made at random given a certain percentage of other-coloured neighbours by all agents just as there will never be a frictionless horizontal plane with a cart with frictionless wheels (as in Galileo's thought experiment). What Galileo's thought experiment and agent-based model have in common, and what makes them good examples of their kind, is that they