Construct Validity and Theoretical Embeddedness of Agent-based Models of Normative Behaviour

Maria Xenitidou, University of Surrey, UK
Corinna Elsenbroich, University of Surrey, UK

Abstract

In this paper we assess the construct validity and theoretical emdeddedness of agent-based models of normative behaviour drawing on experimental social psychology. We contend that social psychology and agent-based modelling share the focus of 'observing' the processes and outcomes of the interaction of individual agents. The paper focuses on two from a taxonomy of agent-based models of normative behaviour. This enables the identification of the assumptions the models are built on and in turn, reflection on the assumptions themselves from a socio-psychological perspective.

Keywords: Agent-based Modelling, Social Psychology, Normative Behaviour, Norms, Construct Validity

1. Introduction

This paper assesses the theoretical embeddedness and construct validity of agent-based models of normative behaviour with respect to experimental social psychology.

Agent-based modelling (ABM) is a computational method used to simulate social phenomena focusing on the interactions of individual agents (see for example Gilbert, 2008). Construct validity is a term usually applied to the measurement of concepts in quantitative social research. In ABM the problem is not principally one of measurement but whether the operationalisations of concepts are appropriate. A theoretical embedding relates the concepts used in ABM to theories in social psychology.

Our work can be seen as an extension to recent work on the construct validity of ABM of normative behaviour (Neumann, 2010). Whilst Neumann focuses on individual psychology and problems of norm internalisation, we focus on social psychology. Although individual psychology is of great importance for the explanation of normative behaviour when it comes to internalisation, social psychology is more appropriate as a theoretical backdrop to ABM.

The need for a theoretical embedding and an assessment of the construct validity of ABM arises from the tricky task of verifying that a model corresponds to its target system. A model is an abstraction from reality. Decisions about which features to keep as salient and which ones to strip are often informed by the use of the model. For example, a model airplane might have the colour and proportions of a real airplane but none of its flying behaviour, whilst a flight simulator for a Boeing 747 simulates the flying behaviour but not the materiality of the real plane. Thus, when

modelling we need to keep in mind what a specific model is for and what a model is of. ABM, like any good science, needs to be validated with appropriate models of target systems. This can be done at different levels of abstraction. For some simulations detailed data are at hand to which a model's behaviour can be compared, but for others the model outcome can only be compared to more general system behaviour. In the study of norms, empirical validation is particularly difficult because phenomena associated with norms (e.g. local conformity, global diversity) are general rather than based on specific data.

We use experimental social psychology as the theoretical backdrop of ABM for three reasons: the first is the procedural similarity which seems to underpin the experimental paradigm in social psychology and ABM both of which test social behaviour (direct or indirect interaction between agents) and its consequences by manipulating various (behavioural and contextual-situational) conditions.

Second, the choice of experimental social psychology is informed by the wide reach that experimental findings had in the 1950s, 1960s and 1970s and the resilience of the experimental approach in social and group influence research to date. Social psychology constitutes the dominant framework for conceptualising normative behaviour, thus inspiring ways to approach the study of norms in the social sciences. Although Critical Social Psychology, which has developed as a critique to experimental social psychology, has emerged as the 'new mainstream' in social psychology (Ibanez, 1997; for an overview of critical social psychology, see *inter alia* Howitt *et al.*, 1989; Wetherell, 1996), experimental social psychology has merit for making connections with ABM on social norms.

Finally, much of ABM work is based on assumptions that can be traced back to experimental social psychology. With regards to the study of social norms in particular, three normative behaviours have been identified in experimental investigations of social and group influence in social psychology: obedience, conformity and compliance. In Section 4 we relate ABM to these concepts of normative behaviour.

The paper is structured as follows. We briefly describe the methodology of agent-based modelling in Section 2, present a tripartite classification of agent-based models of norms and focus on two models in order to discuss theoretical embeddedness and construct validity. Section 3 discusses norms research in social psychology, presenting a classification of normative behaviour in terms of

conformity, obedience and compliance. In Section 4 we position ourselves with regards to these models. We reflect on the two models and point at a shortcoming in ABM in not modelling compliance. Section 5 concludes and points to future work.

2. Agent-based Modelling

Agent-based modelling has been a ground breaking methodological innovation, introducing a new, model-centred epistemology in social science. Narrative methods and/or empirical data alone seem inadequate to capture the complexity of social phenomena. Here complexity refers to multiple, asymmetric, non-linear interactions between autonomous, heterogeneous agents with bounded rationality and incomplete knowledge about the behaviour of the whole system in which they are embedded. The most indicative manifestation of complexity is that simple activities between agents contribute to the emergence of complex social phenomena that cannot be reduced to the behaviours of individual agents (cf. Squazzoni, *forthcoming*). ABM is a methodology that integrates, tests and extends theoretical explanations of social phenomena. It also represents and tries to understand social phenomena grounded in empirical data by capturing their connections and processes. However, the relationship between theory-model-phenomena is not consistently represented in agent-based models (see Squazzoni, *forthcoming*; Epstein, 2006; Henrickson and McKelvey, 2002).

ABM could be considered the `virtual laboratory of the social sciences' as it allows for the controlled variation of parameters, identification of mechanisms and testing of theories similar to laboratory experiments in the natural sciences. Agent-based models are computer programs consisting of a number of autonomous, heterogeneous agents interacting with an environment and with each other. *Autonomy* means that there is no centralised decision maker. Each agent has its own set of plans, rules and parameters determining its behaviour. Agents can be *heterogeneous* in their attributes and preferences as well as in their behavioural rules. For example, in an ABM of the English housing market (Gilbert, 2008), there are agents acting as buyers, sellers and realtors (estate agents) according to different behaviour rules. The agents also have different incomes and savings and are at different stages of the housing ladder (e.g. first time buyers). The *environment* in an agent-based model can represent several different things such as 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. Agents also interact with each other, e.g. exchanging information or

one imitating another. Given these ingredients many macro phenomena can be generated from the interactions of agents at the micro level. ABMs have successfully been employed in the study of opinion dynamics, technology adoption, markets, innovation networks and, our focus in this article, the study of normative behaviour.

2.1 Agent-Based Models of Norms

ABM have their origin in three different research camps; game theory, the natural sciences and distributed artificial intelligence (DAI). Models originating from game theory are studied to examine how cooperation can emerge from the interaction of selfish agents. The search is for payoff structures, incentives and disincentives that produce pro-social behaviour. The most famous examples of the integration of game theory into simulations is Axelrod's work on the emergence and evolution of norms (Axelrod, 1984; 1986).

The strand of modelling coming from the natural sciences uses work on ecology, population dynamics, or particle physics to model human behaviour. The fundamental idea is that there are many complex phenomena in the natural world and that we can use these complexities to analyse the social world. Many cellular automata models such as the game of life (Berlekamp *et al.,* 1982) or zero intelligence models of economic markets (Mirowski, 1999) result from this approach.

The DAI strand starts from a systems perspective. The aim of traditional AI is to find solutions for complex problems by the application of 'intelligence'. Some problems have been recognised as either unsolvable or as being less effectively solved by a single, monolithic system. DAI developed decentralised solutions for some of these problems. Rather than modelling a system as a whole, autonomous component parts are modelled, producing solutions via their interaction (Woolridge, 2002).

These three origins differ markedly in terms of their view, use and appreciation of norms. Agent-based simulations of norms also vary according to which strand they are closer to. We can distinguish three basic kinds of agent-based simulations of norms and roughly attribute them to the three origins as in Table 1.

Table 1: Models of Norms, Aims and Origins

Models of the **Emergence** of Norms	Game theory
Models of the **Diffusion** of Norms	Natural Sciences
Models of the **Use/Application** of Norms	Distributed Artificial Intelligence

It is unclear that the term "norm" differs widely from one field to another. Compare three definitions from the three fields:

> "A norm exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way." (Axelrod, 1986, p. 1097)

> "First, [social norms] are self enforcing behavioural regularities. But second, once entrenched, we conform *without thinking about it*." (Epstein, 2001, p. 1)

> "A normative system is a social construction a society of agents uses to achieve a certain social order." (van der Torre and Boella, 2004, p. 201)

One could say that the three definitions simply focus on different aspects of norms, depending on their particular research question regarding normative behaviour (origin, diffusion or application). However, these definitions appear in the literature without caveats that a specific subdomain of norms is of interest. Along with the definition of norms and normative behaviour come assumptions about the agents and their interactions. These are the assumptions that need to be stated explicitly and, where possible, integrated into theory.

We now discuss two influential models from the ABM literature from which the two first definitions are taken. We shall not discuss a model of the last kind in detail for reasons of space, but shall return to a DAI model briefly in Section 4.

2.2 Emergence of Norms

One of the most famous simulations of norm emergence is Axelrod's (1984; 1986). Based on game theory, agents have a choice between two strategic actions: to cooperate/defect and to punish/not-punish if they observe a defection. The strategy is determined by two parameters: boldness (risk accepted for defection) and vengefulness (likelihood of punishing a defecting agent). By definition cooperation is more costly than defection and being punished is more costly than cooperation. There is also a cost associated with punishing a defector (rather than turning a blind eye).
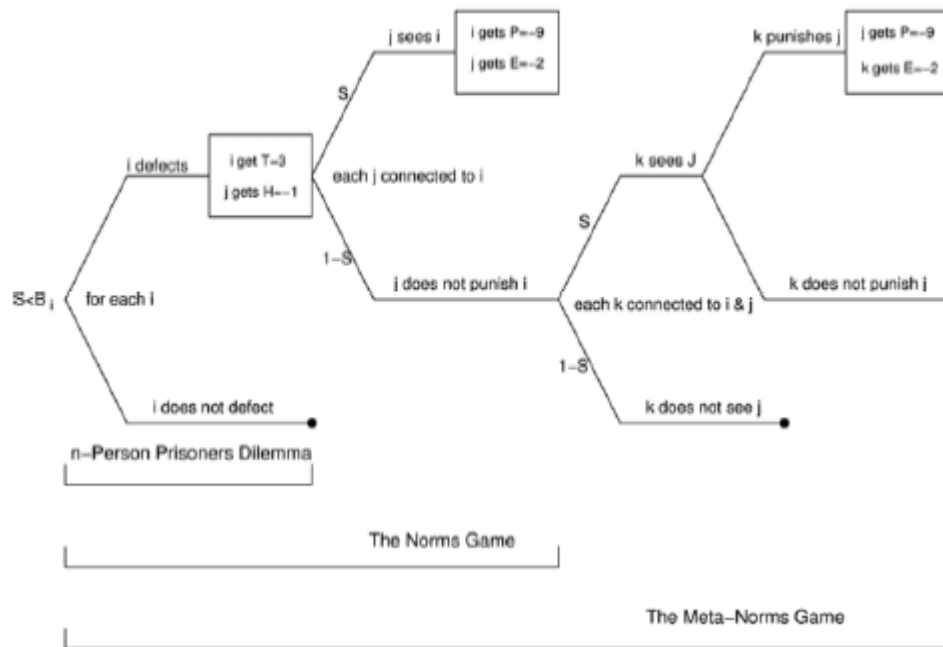

Figure 1: Axelrod's Norm and Meta-Norm Game

Axelrod's original experiments define agents that play 1000 rounds of prisoners' dilemma games. When the norm game is played, i.e. defectors are punished at the cost of the punisher, the outcome of the simulation experiments is that norms ultimately collapse. The vengefulness of the population initially reduces defection but after a period of norm adherence, vengefulness levels fall, leading to a rise in defection and irretrievable norm breakdown. Axelrod then adds a meta-norm game, in which an agent that is observed not punishing a defector is punished. In this simulation, cooperation and thus a stable norm emerges, under a wide range of initial conditions. There has been some

criticism of Axelrod's simulations, showing that the meta-norm game also breaks down eventually and that a much larger number of agents produce relevantly different results (Mahmoud, 2010).

The meta-norm simulation assumes: a) Hobbesian view of human nature as selfish and needing to be constrained, b) norms are necessarily costly for the individual, c) norms are sustained by punishment, and d) the decision to cooperate or defect is a simple calculation and not dependent on the social setting (that is, there is no influence from other agents except for the possibility of punishment).

2.3 Norm Adoption and Diffusion

Models concerned with norm adoption and diffusion do not ask how general normative behaviour emerges but how a specific norm might be chosen and then diffuse through a population. There are several different kinds of diffusion model. Some are based on social learning (e.g. Lorscheid, 2009), some on memetics and imitation (e.g. Gatherer, 2002, Flentge, 2001) and some on opinion dynamics (e.g. Deffuant *et al.*, 2002). Both imitation and opinion dynamics models have their theoretical foundations in natural science dynamic models rather than the social sciences. Social learning models are a recent development relying on cognitively more complex agents than traditional ABM.[1]

As an example of a model of diffusion, we focus on one described by Epstein (2000). Epstein states two macro phenomena of normative behaviour: a) norms are locally stable, i.e. a group of agents close to each other adhere to the same norms and b) norms are globally diverse. A micro assumption of the model is that agents adhere to norms "without thinking about them". Epstein operationalises this assumption as varying computing power of the agents. The model starts by assigning a binary norm and a norm change mechanism to agents. Norm change depends on two variables: the local environment and the amount of thinking about behaviour an agent does. At each step, every agent checks the norm of its neighbours within a personal vision radius. If the agent is in accord with its surrounding it keeps its norm and decreases its vision radius (less thinking). If it is not in accord it changes the norm and increases its vision radius (more thinking). The model's results show the desired macro-behaviour, leading to local conformity and global diversity of norms. Agents at the heart of a norm no longer think about their behaviour whereas those at the

---

[1] See Section 4 for a brief discussion.

borderline between norms have to choose between competing local norms. The model also captures the re-emergence of norm differentiation after a 'revolution', a shock to the system.

The model assumes: a) the automatic adoption of the majority norm, b) the non-thinking when agents are in a stable norm-environment, c) the non-enforcement of norms.

These two models, one drawing from game theory and the other from statistical physics, will serve as examples for the following discussion of the agent-based modelling approach to understanding normative behaviour. In the next section, we consider how social psychology has treated explaining normative behaviour.

3. Social Psychology

Social psychology departed from psychology in treating the 'social' as a key element of psychological processes. According to social psychology, psychological processes are both social and cognitive. These are premises shared by agent-based modelling with its focus on agent interaction. Social processes are the ways in which perceptions, thoughts, emotions, memories, motives and actions are influenced by other people such as friends, personal relationships or family. Cognitive processes are the ways in which perceptions, thoughts and memories guide our understanding of the world.

3.1 Social Psychological Approaches to Norms

The study of norms and normative behaviour in social psychology is a branch of social influence research – research into the ways people affect the thoughts, feelings and behaviours of others. Social cognition perspectives on social influence, using experimental (and, later, field) methods, have dominated the study of norms. The classic experimental studies lead to a tripartite classification into *conformity, obedience* and *compliance*.

Conformity refers to yielding to perceived group pressure by copying the behaviour and beliefs of others (Franzoi, 2000). An agent conforms if it simply follows the behaviour of other agents. An agent can conform to a majority or a subgroup of the population. We drive on the left hand side of the road when in Great Britain but on the right on the European continent. We might wear bumbags and socks when on holiday in Italy just like other tourists or wear sunglasses and black like the Italians do. In either case, we conform to a subset of the population.

Conformity research has its foundations in Sherif's (1936) Norm Development Research, which consisted of lab experiments on optical illusions (the Autokinetic Effect) without prior information or a frame of reference. The conclusion drawn was that where reality is ambiguous and fluid, people develop and internalise shared 'frames of reference' to introduce order, stability and coherence into their relations with the stimulus world and each other (Sherif, 1936; see discussion in Turner, 1991). This line of research emphasised the development of norms as a consequence of *informational* influence that is, the desire to obtain useful/accurate information and, thus, achieve a correct or appropriate result.

Asch (1952) disagreed with the idea that group influence is arbitrary and hypothesised that when reality is not ambiguous and individuals are presented with 'facts', they would not conform to a developing consensual norm but will remain independent. In his Line Judgment Experiments he showed (in one third of the trials) that norms are not shared frames of reference but group evaluations of what is right and wrong. Conforming to these norms involves converging individual perceptions and behaviour to that of the group. In other words, Asch's work underlined the element of group pressure in the formation and conformity to a norm over that of consensual and unanimous formation and conformity (Sherif, 1936). Thus, this line of research emphasised the development and conformity to norms as a consequence of *normative* influence, the tendency to conform in order to gain social acceptance and avoid social rejection or conflict.

More recent research has concluded that over time social norms become associated with specific settings. The situation (stimuli) 'activates' remembered mental representations of normative behaviours (e.g. Stereotype Activation) (Dijksterhuis and Bargh, 2001; Aarts and Dijksterhuis, 2003). These are also called 'situational' norms and are activated in anticipation of entering a particular – already experienced – situation, e.g. being quiet when entering a place of prayer or a library. This research stressed how social norms become diffused and thus normative behaviour itself becomes routine, being executed without involving conscious thought.

Obedience refers to the performance of an action in response to a direct order or command. An agent obeys a norm if there is an authority enforcing the norm, e.g. by punishment. The punishment can be applied by a small group or even a single person with authority, for example, the police or a father. Punishment can also be applied by ostracism by the agent's community. We do not jump a red light when driving because we are afraid of being caught and fined by the police.

The classic research on obedience is that of Milgram (1974), which consisted of lab experiments involving subjects role-playing 'teachers' and 'learners'. Physical punishment in the form of electric shocks was administered by teachers to learners when instructed by the experimenter. The 'learners' played an act of being shocked but did not receive real shocks. The experiments indicated the pervasiveness of social power and status in situations of obedience. Authority, such as playing the role of experimenter, increases the likelihood that commands will be obeyed.

Obedience is treated as a demand on the agent to make a binary choice. If the external pressure becomes too big and is exercised by a more 'powerful' agent, the agent will change its behaviour and obey the norm. The agent needs to understand the cost of defying a norm, i.e. the possibility of punishment. Nothing seems to change in the agent between the pre-norm adoption and post-norm adoption state (see Milgram, 1974; cf. Gibson and Haritos-Fatouros, 1986). Norm adherence is fully determined externally. Analyses of norms that assume that agents follow the mechanism of obedience are theories based on a Hobbesian conception of society constraining the individual, such as Rational Choice theory and game theory (e.g. Coleman, 1990; Bichieri, 2006). These theories, relying on individualism, have been accused of not taking the social influence on agents seriously.

Compliance refers to acting publicly in accord with a direct request or suggestion. Compliance is often treated as deliberate social influence (Hewstone *et al.*, 2008) in the sense that individuals make conscious and deliberate attempts to gain the social approval of others. Factors that foster compliance are *positive moods* (likelihood of compliance when in good mood), *reciprocity* (the expectation that one should return a favour or a good deed), and *giving reasons* (explanations that make requests reasonable. All come with the precondition of equal social status amongst the individuals who interact, as otherwise compliance might reduce to obedience (Franzoi, 2000; cf. Moscovici, 1976). Compliance entails an assessment of the situation, thus individuals behave as participants rather than recipients in social encounters. An agent complies with a norm if it follows it while being able to consider other behaviours. To comply contains the element of choice, that is, to comply with something now but not later or to comply with something superficially but 'not really'. As compliance is associated with publicly acting in a certain way whilst retaining one's personal beliefs, we might comply with wearing a suit to work although we would never wear formal clothes otherwise.

In experimental terms, compliance is a more complex form of norm adherence than the other two forms. Although in conformity the agent's mental change is complete and in obedience there is no

change at all, compliance involves the possibility of gradual change. The concession of the individual to social pressure while attempting to maintain their individual values, may have the effect of gradually changing their values. Compliance demands representations of the world 'given compliance' and the world 'without compliance' and involves agents making behaviour choices that are sensitive to the context. It also demands intentionality; the agent deliberately decides whether to comply with a norm or not. The interplay of these representations and the intentionality and context-sensitivity of behaviour choices is where the interesting feedback between the individual and society lies.

4. Reflections on Experimental Social Psychology and Agent-based Modelling

Examining the assumptions of the two agent-based models of normative behaviour discussed above, we find that Axelrod's model is based on the obedience paradigm and Epstein's on the conformity paradigm. If we go back to Axelrod's and Epstein's definitions of norms quoted above, we see that Axelrod's definition of a norm as something that "exists in a given social setting to the extent that individuals usually act in a certain way and are often punished when seen not to be acting in this way" relates to obedience through the punishment aspect, although it leaves out other salient features of obedience such as authority and social hierarchy. Epstein's conceptualisation of norms as "self enforcing behavioural regularities […] entrenched without thinking about them" is very close to the recent research on thoughtless conformity (Dijksterhuis and Bargh, 2001; Aarts and Dijksterhuis, 2003).

Considering these two models through the lens of experimental social psychology's notions of norms highlights some failings in the ways that the models implement normative behaviour. For instance, treating conformity as behaviour implying thoughtlessness has been a matter for criticism of mainstream experimental social psychology (see Howitt *et al.*, 1989; Wetherell, 1996; Lesko, 2003) and these criticisms can be extended to the agent-based models. Moreover, the models do not even fulfil their potential of representing the essentially dynamic nature of the development of norms. For example, conformity is a gradual process taking a finite time, while both models implement conformity as an instantaneous switch from one behaviour to another. Systematic exposure to obedience structures may be needed to induce conformity, particularly in institutionalised settings (e.g. training a torturer, Gibson and Haritos-Fatouros, 1986; 1971 Stanford Prison Experiment, Zimbardo, 2007).

In addition, identities and roles need to be taken into account. Thus rather than treating behaviour as determined by the individual agent's state plus the situation, the different meanings that a situation may have for different agents in different roles and with different histories need to be taken into account. If one does so, behavioural consequences might be expected to emerge, such as varieties of *non-conformity* – not doing what is expected owing to personal, situational, cultural factors (see Franzoi, 2000) – and *minority influence* – a faction within a group influencing the group during decision making (Moscovici, 1976). The type of norm is another factor to be considered when studying normative behaviour. For example, it may be easier to conform to perceptions of the physical world or to driving on the right than to particular social values (Smith, 1998).

Agent-based models of normative behaviour undermine the 'social' and do not embed agents in the social processes they are part of. Axelrod's model takes the 'social' out of the mix by modelling individual responses to punishment. Epstein's model takes the 'social' out because there is nothing distinctly specific to human agents in how norms spread in his model. In this way, a division between the individual and the social, and between agency and structure is maintained. Social influence processes are social practices and social interactions and should be modelled as an agent interaction outcome if we wish to do better than produce explanations that reproduce the individual and the social, agency and structure as distinct entities.

Although, we have discussed two models, we have presented three approaches of norm adherence in experimental social psychology (conformity, obedience and compliance). The reason why we did not discuss a specific model exemplifying compliance is that we believe there is, as yet, no such model. It is to models based on distributed artificial intelligence that we should look to find models based on compliance, because compliance requires that we model cognition, and only DAI models offer this possibility. The current model that comes closest to representing compliance is the agent architecture EMIL-A (Andrighetto *et al.*, 2007) and the corresponding simulations. This architecture enables the agent to adopt new norms from the outside world and integrate them into its decision procedure. For example, the EMIL simulation of normative learning incorporates reinforcement learning, social approval and sanction into the agent's decision-making (Lorscheid, 2009). A particularly interesting aspect is what the authors call the *penitence level* of an agent. Penitence is a heterogeneous agent variable (ranging from embarrassed, through guilty, to impervious). Depending on the penitence level, agents can resist changing their internal values in the face of more or less negative norm invocation messages. Thus, although the agent now has the

possibility of keeping its initial norms for a time and slowly adopting new norms, the mechanism underlying it is a sort of punishment formalised in the norm invocation messages and the penitence level. An extension to the architecture might be able to display the context sensitivity necessary for norm compliance.

5. Conclusions and Future Work

We have shown that a theoretical embedding of ABM in concepts drawn from experimental social psychology can be useful for the validation of normative models. Current models focus on one kind of norm adherence only, emergence models on obedience and diffusion models on conformity. Compliance is not yet tackled by ABM, although there is recent work going in that direction (e.g. Andrighetto *et al.*, 2007) and there are rich opportunities for developing this approach. For example, modelling compliance would entail simulating an individual's negotiation of the personal (desires, beliefs), the interpersonal (the meaning of the situation, social roles, norms, and the meaning of norms) and the experiential (memory) (drawing on Howitt *et al.*, 1989).

We are looking to develop the work on compliance in more depth to lay open the challenges it poses to ABM, in particular the need for modelling the cognition of an agent. In this paper, we have focussed only on experimental social psychology. In future research, we shall also consider the contributions of other strands of social psychology. For example *critical social psychology* has a dynamic view of norms, something in tune with ABM, given that the normative behaviour can be generated dynamically by modelling.

Acknowledgement

References

Aarts, H., and Dijksterhuis, A. (2003) The silence of the library: Environmental control over social behavior. *Journal of Personality and Social Psychology*. 84, pp. 18-28.

Andrighetto, G., Conte, R., and Turrini, P. (2007) *Emergence in the loop: Simulating the two-way dynamics of norm innovation*. In Dagstuhl Seminar Proceedings 07122 Normative Multi-agent Systems. http://drops.dagstuhl.de/opus/volltexte/2007/90707122, D. S. P., editor, Normative Multi-agent Systems.

Aronson, E. (1999) *The Social Animal*. New York: Worth/Freeman.

Asch, S. E. (1952) *Social Psychology*. New Jersey: Prentice Hall.

Asch, S. E. (1956) Studies on independence and conformity. *Psychological Monographs*. 70(a), pp. 1-70.

Axelrod, R. (1984) *The evolution of cooperation*. Basic Books: New York.

Axelrod, R. (1986) An evolutionary approach to norms. *American Political Science Review*. 80, pp. 1095-1111.

Berlekamp, E., Conway, J. and Guy, R. (1982) *Winning Ways for Your Mathematical Plays. Vol.: 2 Games in particular*. London: Academic Press.

Bichieri, C. (2006) *The Grammar of Society: The Nature and Dynamics of Social Norms*. New York: Cambridge University Press.

Boella, G. and van der Torre, L. (2004) *An Agent Oriented Ontology of Social Reality*. In Proceedings of Formal Ontologies in Information Systems conference (FOIS'04), Amsterdam, pp. 199-209. IOS Press.

Burke, M. and Fournier, G. (2005) The Emergence of Local Norms in Networks. *Computing in Economics and Finance* 2005 299, Society for Computational Economics.

Conte, R. and Castelfranci, C. (1995) Understanding the functions of norms in social groups through simulation. In N. Gilbert and R. Conte (Eds.) *Artificial Societies. The computer simulation of social life*. London: UCL Press.

Castelfranchi, C., Conte, R. and Paolucci, M. (1998) Normative reputation and the cost of compliance. *Journal of Artificial Societies and Social Simulation*. 1(3).

Deffuant, G., Amblard, F., Weisbuch, G. and Faure, T. (2002) How can extremism prevail? A study based on the relative agreement interaction model. *Journal of Artificial Societies and Social Simulation*. Vol. 5, No. 4.

Dignum, F., Kinney, D. and Sonenberg, L. (2002). From desires, obligations and norms to goals. *Cognitive Science Quarterly*. 2(3/4).

Dijksterhuis, A., and Bargh, J. A. (2001) The perception-behavior Expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology*. 33, pp. 1-40.

Epstein, J. and Axtell, R. (1996) *Growing Artificial Societies: Social Science From The Bottom Up*. Washington: Brookings Institute Press.

Epstein, J. (2000) *Learning to be thoughtless: social norms and individual computing*. Center on Social and Economic Dynamics. Working Paper, No. 6.

Epstein, J. (2006) *Generative Social Science. Studies in agent-based computational modeling*. New Jersey: Princeton University Press.

Flentge, F., Polani, D. and Uthmann, T. (2001) Modelling the emergence of possession norms using memes. *Journal of Artificial Societies and Social Simulation*. 4(4).

Franzoi, S. L. (2000) *Social Psychology* (2nd ed.). New York: Brown and Benchmark.

Gatherer, D. (2002) Identifying cases of social contagion using memetic isolation: comparison of the dynamics of a multisociety simulation with an ethnographic data set. *Journal of Artificial Societies and Social Simulation*. 5(4):5.

Gibson, J. T. and Haritos-Fatouros, M. (1986) The Education of a Torturer. *Psychology Today*. No. 20, pp. 50-58.

Gilbert, N. (2004) *Agent-based social simulation: dealing with complexity*.

http://cress.soc.surrey.ac.uk/web/resources/ABSS%20%20dealing%20with%20complexity

1.pdf.

Gilbert, N. (2007) *Agent-based models*. London: SAGE.

Gilbert, N. and Troitzsch, K. G. (2005) *Simulation for the social scientist* (2nd ed.). Milton Keynes: Open University Press.

Gilbert, N., Hawksworth, J. C., and Swinney, P. A. (2008) *An agent- based model of the UK housing market*. Technical report, CRESS, University of Surrey, http://cress.soc.surrey.ac.uk/housingmarket/ukhm.html.

Henrickson, L. and McKelvey, B. (2002) Foundations of "new" social science: Institutional legitimacy from philosophy, complexity science, postmodernism, and agent-based modelling. Proceedings of the National Academy of Sciences. 14 May 2002, pp. 7288-7295.

Hepburn, A. (2003) *An Introduction to Critical Social Psychology*. London: SAGE.

Hewstone, M., Stroebe, W. and Jonas, K. (Eds.) (2008) *Introduction to social psychology: a European perspective*. (4th ed.) Oxford: Blackwell.

Howitt, D., Billig, M., Cramer, D., Edwards, D., Kniveton, B., Potter, J. and Radley, A. (1989) *Social Psychology: Conflicts and Continuities*. Milton Keynes and Philadelphia, PA: Open University Press.

Ibanez, T. and Iniguez, L. (1997) (Eds.) *Critical Social Psychology*. London: SAGE.

Lesko, A. .W. (Ed.) (2003) *Readings in Social Psychology: General, Classic, and Contemporary Selections*. (5th ed). Boston: Pearson Education, Inc.

Lorscheid, I. and Troitzsch, K. G. (2009) .*How do agents learn to behave normatively? Machine learning concepts for norm learning in the EMIL project*. In Proceedings of the 6th Annual Conference of the European Social Simulation Association, Guildford, UK, September 2009.

Mahmoud, S., Griffiths, N., Keppens, J. and Luck, M. (2010) An Analysis of Norm Emergence in Axelrod's Model. *NorMAS*.

Milgram, S. (1974) The Perils of Obedience. *Harper's Magazine.*

Mirowski, P. (1999) *More Heat than Light: Economics as Social Physics, Physics as Nature's Economics*. Cambridge, UK: Cambridge University Press.

Moscovici, S. (2000) *Social Representations. Explorations in Social Psychology*. Polity Press in association with Blackwell Publishers.

Potter, J. (1996) Attitudes, Social Representations and Discursive Psychology. In M. Wetherell (Ed.) *Identities, Groups and Social Issues*. London: SAGE, pp. 119-173.

Saam, N. J. and Harrer, A. (1999) Simulating norms, social inequality, and functional change in artificial societies, Journal of Artificial Societies and Social Simulation 2.

Sherif, M. (1936) *The Psychology of Social Norms*. New York: Harper and Raw Publishers.

Smith, P. B. (1988) Norms and roles in the small group. In G. M. Breakwell, H. Foot and R. Gilmour (Eds.) *Doing Social Psychology. Laboratory and Field Exercises*. New York: Cambridge University Press.

Squazzoni, F. (Forthcoming) The Impact of Agent-Based Models in the Social Sciences After 15 Years of Incursions. *History of Economic Ideas.*

Stainton Rogers, W. (2003) *Social Psychology: Experimental and Critical Approaches*. McGrawHill Education.

Turner, J. C. (1991) *Social Influence*. Milton Keynes: Open University Press.

Wetherell, M. (Ed.) (1996) *Identities, groups and social issues*. Milton Keynes: Open University Press.

Wooldridge, M. (2002) *An Introduction to MultiAgent Systems*. Chichester: John Wiley and Sons Ltd.

Zimbardo, P. G. (2007) *The Lucifer Effect: Understanding how good people turn evil*. New York: Random House.