# City Research Online

## City St George's, University of London

**Citation:** Baronchelli, A. (2018). The Emergence of Consensus: a primer. Royal Society Open Science, 5(2), 172189. doi: 10.1098/rsos.172189

This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

**Permanent repository link:** https://openaccess.city.ac.uk/id/eprint/17655/

**Link to published version:** https://doi.org/10.1098/rsos.172189

**Copyright and Reuse:** Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to City Research Online policy.

City Research Online: http://openaccess.city.ac.uk/ publications@citystgeorges.ac.uk

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018

ROYAL SOCIETY OPEN SCIENCE

rsos.royalsocietypublishing.org

Review

**Cite this article:** Baronchelli A. 2018 The emergence of consensus: a primer. *R. Soc. open sci.* **5**: 172189.
http://dx.doi.org/10.1098/rsos.172189

Received: 12 December 2017
Accepted: 18 January 2018

**Subject Category:**
Physics

**Subject Areas:**
complexity/computer modelling and simulation

**Keywords:**
social consensus, spontaneous ordering, computational social science, network science, modelling, experiments

**Author for correspondence:**
Andrea Baronchelli
e-mail: andrea.baronchelli.1@city.ac.uk

THE ROYAL SOCIETY PUBLISHING

# The emergence of consensus: a primer

Andrea Baronchelli

Department of Mathematics, City, University of London, Northampton Square, London EC1V 0HB, UK

AB, 0000-0002-0255-0829

The origin of population-scale coordination has puzzled philosophers and scientists for centuries. Recently, game theory, evolutionary approaches and complex systems science have provided quantitative insights on the mechanisms of social consensus. However, the literature is vast and widely scattered across fields, making it hard for the single researcher to navigate it. This short review aims to provide a compact overview of the main dimensions over which the debate has unfolded and to discuss some representative examples. It focuses on those situations in which consensus emerges 'spontaneously' in the absence of centralized institutions and covers topics that include the macroscopic consequences of the different microscopic rules of behavioural contagion, the role of social networks and the mechanisms that prevent the formation of a consensus or alter it after it has emerged. Special attention is devoted to the recent wave of experiments on the emergence of consensus in social systems.

## 1. Introduction

Money, language, dress codes, decorum, notions of fairness all need to be accepted and shared at the group level in order to function. They require social consensus and in exchange they provide individuals with expectations on how others will behave, eventually allowing a society to operate [1,2]. But how does consensus (or 'order', 'coordination', 'agreement') emerge out of an initially disordered situation when there is more than one possible equilibrium?

This question is key to the Social Sciences and to a wide array of disciplines, ranging from Biology to Physics and from Ethology to Artificial Intelligence. In fact, nature offers countless examples of initially disordered collections of agents that are able to develop shared coordinated behaviours. Flocks of birds frequently change their flight direction [3], fish schools display spontaneous evasion waves [4], ferromagnetism is the result of ordering spins [5] and designing decentralized artificial systems is one of the frontiers of Robotics [6]. All these areas have contributed to advance our comprehension of the mechanisms of consensus [2,7–10].

This interdisciplinary interest has determined the recent explosion in the number of scientific articles investigating the

© 2018 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





emergence of consensus, with two consequences. On the one hand, the similarity between explanations proposed in different areas risks going unnoticed due to different jargons and problem-specific details. On the other hand, communities of researchers exist that largely ignore each other even within apparently confined contexts. For example, two recent and insightful reviews mainly concerned with the problem of consensus, 'The Evolution of Norms' [9] and 'The Evolution of Social Norms' [10], do not share a single bibliographic entry. At the same time, in the last few years it has become clear that understanding the interplay between social consensus and our collective behaviour is crucial to address many of the issues faced by our complex society such as climate change, biodiversity loss and antibiotic resistance [11].

A systematic review of the literature on the emergence of consensus is out of the scope of the present paper, which offers a very brief introduction to the subject. My aim is to provide an overview of the most important principles over which the debate on social consensus has unfolded, and to discuss their implications with the aid of a few *illustrative examples*. By adopting the language of social conventions, possibly the simplest example of social consensus taken from the Social Sciences (§2), I will start by mapping the landscape of proposed solutions to the problem of consensus (§3) before focusing on the case of spontaneous emergence in the absence of a centralized authority (§4). In this context, by considering two simple models, I will discuss how different kinds of behavioural contagion and social networks influence the dynamics of collective agreement (§5), as well as which mechanisms can either alter (§6), or hinder or prevent consensus (§7). Finally, I will provide an overview of recent experiments that provide empirical basis to the study of the emergence of consensus in social systems (§8).

## 2. A prototypical example: social conventions

The word 'spam' refers to 'disruptive online messages [...] sent as email' [12]. However, the Internet and—as a consequence—the phenomenon that today we indicate as spam did not exist just a few decades ago. So how did we end up agreeing that those annoying messages are to be called 'spam'? Or, actually, how did we *manage to* agree?

Naming conventions have attracted the attention of philosophers since the ancient past. Hermogenes, in Plato's Cratylus, asserts that names belong to things 'only because of the rules and usages of those who establish the usage and call it by that name' [13], without commenting on *how* a group reaches consensus on a specific name. On the other hand, Adam, the first human in the Bible, establishes new names for the objects around him [14]. Far from being curiosities, these two solutions identify a first major divide between different approaches. Consensus can be imposed by an authority or emerge from an interacting multitude.

Conventions govern much of social and economic life. In general, a convention is a pattern of behaviour that is customary, expected and self-enforcing [1,15]. It is the result of a coordination process where one among various different alternatives is adopted, and they are maintained because a unilateral deviation makes everyone worse off [15]. Among the countless situations where consensus plays a crucial role, some of which have been mentioned above, this review adopts the perspective—and language—of social conventions both for its transparency for readers with different backgrounds and for its historical prominence.

## 3. Modelling consensus

The emergence of consensus can be described both as a *cooperative process* in the space of individuals trying to coordinate with each other, and as a *competitive process* in the space of the alternatives individuals can adopt. Different approaches make different hypotheses on the structure of these two spaces. Here, we will consider only theories that describe consensus as the result of the interactions between individuals [16,17], but different possibilities exist [13,18,19].

### 3.1. Space of individuals

A population can be described in terms of a network whose nodes represent individuals and links identify potential interactions. The coordination between groups of neighbouring nodes is referred to as *local* consensus, while *global* consensus indicates that (most of) the population has reached an agreement. The structure of the social network plays a major role on the dynamics of consensus, in ways that depend on the details of the microscopic individual interactions (see §5). A major distinction between different

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





models of consensus concerns the presence and role of a formal or informal centralized institution, or in general of any actors or mechanisms able to exert a global influence on the system.

(i) Prominent examples of centralized institutions are:
   - *Authority*. An authority that has the means to enforce order through violent or non-violent punishment of the violators is the simplest source of social order [20].
   - *Leadership*. Leaders need to be identified as such based on some merit [21]. Potential leaders include 'connectors', who have a large social circle, 'mavens', who rely on a deep knowledge of a specific topic, and 'persuaders', who have exceptional negotiation skills [22].
   - *Broadcasting*. One to many distributors of information can influence consensus both on a specific opinion or by 'setting the agenda' on a set of acceptable or urgent problems [23,24].
   - *Explicit incentives for collective coordination*. A centralized institution makes individuals aware that they will benefit from global consensus, potentially making them more prone to seek coordination also outside of their immediate social circle [25,26].
   - *Informational feedback*. While no incentives for global coordination exists in this case, individuals are informed about the population-level popularity of the different options [27]. Conformity and social pressure can then favour final consensus [28].

(ii) When a centralized institution does not exist, consensus comes either from the interaction between agents or from some pre-defined individual behaviour. Examples of the two cases are:
   - *Spontaneous emergence of consensus*. Consensus is said to be 'spontaneous' when a centralized institution is not present and agreement is produced by self-interested individuals who are not intentionally aiming at global coordination [29]. The dynamics of the process, or 'evolutionary' forces [10], select the equilibrium [1,29–34]. Important mechanisms that can foster spontaneous consensus are [35]
     - *Communication*. For example, earlier participants can explain the benefits of coordination to latecomers [36], or individuals can negotiate some form of local consensus [37,38].
     - *Peer punishment of deviants*. When the benefit of (local, at least) consensus are greater than the individual cost of punishing her peers or if the cost of being punished is large enough, then sanctions on deviants are a powerful tool to promote consensus [39].
     - *Positive payoff externalities*. This is the case of self-enforcing norms, such as for example driving on the left or on the right of the road. Once established they persist indefinitely [40–45].
     - *Conformity bias*. An inherent tendency to conform to the behaviour of others is a hallmark of human culture [28,46,47] and has been observed also among chimpanzees [48].
   - *Quorum sensing*. Individuals are capable of assessing the number of peers they interact with and share a pre-defined response once a threshold number of components is detected. Bacteria [49], ants [50] and honeybees [51] are examples of social species that use quorum sensing.

## 3.2. Space of alternatives: equilibrium selection

A natural question is which alternative, or equilibrium, will be selected by the population in the case of consensus. Three notable answers are:

- Individuals select a given alternative by logical reflection. They are able to assess the advantages of one equilibrium over the others and act to maximize their benefit. Rational considerations would therefore guide individual choice [52,53].
- Individuals select a given alternative based on psychological, even though not rational, factors. Shared biases select the best alternative to be played [1,40].
- Alternatives are equivalent, and the dynamics of the process where learning individuals interact eventually selects one of the possible equilibria 'by chance' [7,15,32].

Note that only in the latter case communication, or interaction, between individuals is necessary to reach a consensus, as we will see in §4. In the other two cases, in fact, individuals independently select

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





the same alternative based on internal processes [1]. A further dimension affecting the three scenarios concerns the *basin of attraction* of the different alternatives, i.e. the region of the phase space such that any point (any initial condition) in that region will eventually be iterated into the attractor [54,55].

## 4. Approaches to the study of spontaneous consensus

This and the following sections focus on the case of spontaneous emergence of consensus, where the aim is to understand the macroscopic consequences of microscopic behaviours [8,9]. Two main approaches to investigate spontaneous consensus are game theory and the evolutionary—or dynamic—approach.

In coordination games with multiple pure-strategy Nash equilibria, consensus emerges when one equilibrium is selected by all the members of the population. However, it was soon realized that traditional game theory fails to explain how players would know that a Nash equilibrium is to be played and which Nash equilibrium is to be selected when more than one equivalent choices are present [56]. A possible solution is attributing the equilibrium selection at the level of individual decision-making [52], but this requires strong and unrealistic assumptions on the individual access to, and processing of, information [57,58].

Evolutionary explanations overcome this difficulty by substituting actors' rationality and knowledgeability with the capability of anticipating what others will do, and by specifying how individuals learn from experience and adjust their choices accordingly. At least two main frameworks implement this approach. On the one hand, in Evolutionary Game Theory [59] individuals, who are born with a behavioural strategy, interact and reproduce according to a fitness proportional to the payoff of the game they play. Evolution determines over time the successful strategies, possibly driving the population to an equilibrium. Crucially, the biological framing of genetically encoded strategies and reproduction can be translated in terms of bounded rationality and learning when describing social systems [10,31,60]. On the other hand, agent-based modelling aims to understand the global consequences of individual adaptive behaviour relying on the concepts of emergence and self-organization developed in statistical physics. Pioneered by celebrated examples such as Schelling's segregation model [61], Axelrod's work on competition and collaboration [62] and Reynold's flocking model [63], agent-based models have witnessed an explosion in the last two decades thanks to computational methods and numerical simulations, acquiring a central role in the study of social dynamics. A review of these models is out of the scope of the present paper, which is condemned to leave out many interesting contributions, but the interested reader can find an extensive survey in Castellano *et al*. [8].

## 5. Microscopic interactions, social networks and the dynamics of consensus

Multi-agent models define agents that can assume different states, and rules that determine how these states change, typically through interactions. A major distinction concerns the number of times an agent needs to be exposed to another state before adopting it. In simple contagion models, one exposure to a different state may be sufficient [64]. In complex contagion models, on the other hand, more exposures are required, typically from more than one source (if interactions reveal the identity of the individuals) [65]. The consequences of the adopted kind of contagion can be profound and it is useful to see it in two simple models, chosen purely as illustrative examples.

The Moran process was introduced to study selection in a finite population [66]. Individuals are characterized by a state variable that can assume one of $M$ values. In each time step, two neighbouring individuals are randomly chosen, one for reproduction and one for elimination. The offspring of the first individual will replace the second. Equivalently, it can be said that the second individual will adopt the state of the first one, in a process of *simple contagion*. The same dynamics was introduced a second time under the name of voter model [67,68]. Here, the first individual adopts, or 'copies', the state of the second one. The two variants are equivalent on homogeneous topologies but exhibit different behaviours on heterogeneous networks [69–72].

The naming game addresses the emergence of simple (linguistic) conventions following a scheme devised by Wittgenstein [32,34,37], which is very similar to the signalling game introduced by Lewis [1] when decisions based on common knowledge are replaced by adaptive behaviour [73]. Nowadays in the standard formulation [34], individuals are characterized by an inventory of names, which is empty at the beginning of the process. In each time step, a pair of neighbouring agents is chosen randomly, one to play as hearer and the other as speaker. The speaker randomly selects one of its names, or invents a new name if its inventory is empty. If the hearer's inventory contains such a name, the two individuals update their

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





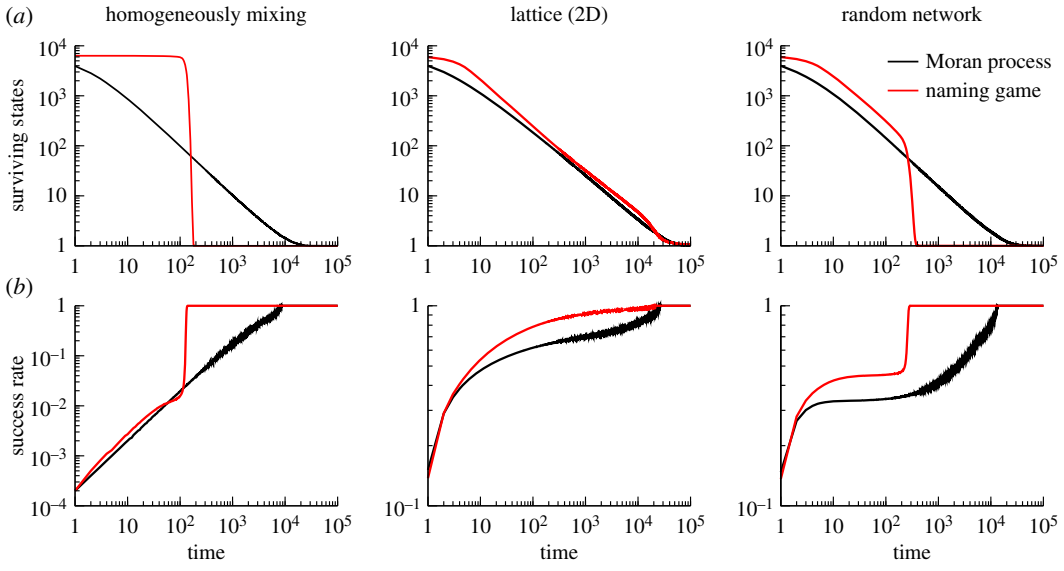


**Figure 1.** Different paths to spontaneous consensus. (*a*) Surviving states for the Moran process (simple contagion) and naming game (complex contagion) on different topologies. (*b*) Success rate, defined as the probability of observing an interaction involving two identical individuals in the Moran process or a successful interaction in the naming game (similar alternative observables exist for the two models, the qualitative description is not affected by the particular choice). In homogeneously mixing populations, the Moran process evolves through a progressive elimination of different states, while the naming game exhibits a sharp transition to order (symmetry breaking). The dynamics of the two models appear more similar on lattices, although profound differences exist (figure 2). On complex networks, on the other hand, after an initial phase in which the two models appear similar, the naming game exhibits a transition to order similar to the one observed on homogeneously mixing populations. Population size of $N = 10\,000$ individuals prepared initially in $M = N$ different states. Lattice and random network have coordination number $k = 4$ for all the nodes.

inventories so as to keep only the word involved in the interaction, otherwise the hearer adds the name to those already stored in its inventory. Thus, at least two interactions are needed for an individual to go from state $A$ to state $B$, a characteristic feature of *complex contagion*. The number of names can be fixed by endowing agents with a name at the beginning of the game.

In finite-size populations, consensus emerges both in the Moran process and naming game and in both cases once it is reached it will persist indefinitely. However, the mechanisms controlling how the population 'selects' the alternative to agree upon are qualitatively different in the two models. To see this, it is convenient to consider different interaction topologies separately. The appendix contains a glossary of network terms.

Before proceeding, it is worth noting that beyond the number of exposures necessary for an agent to change state other factors play an important role. According to the theory of social impact [74], for example, the impact of a group on an individual is proportional to the 'strength' of the members of the group (how credible or persuasive they are), their 'immediacy' (a decreasing function of their social 'distance' from the individual) and their number, $N$. Various multi-agent models explored [75–77] or took inspiration from [78] this approach but unfortunately we are forced to limit our analysis to the above-mentioned examples due to space limitations.

## 5.1. Homogeneously mixing populations

In the Moran process, interactions are symmetrical. If only two states are available, when an agent in state 0 and an agent in state 1 interact, the outcome is either two agents in 0 or two agents in 1 with the same probability ($p = \frac{1}{2}$). Thus, a chain of interactions favouring one state (i.e. a large *fluctuation*) is needed in order for that state to prevail. When more states are available, consensus is reached through a progressive elimination of alternatives (figure 1). In particular, the probability that consensus is reached on state 1 when there are $i$ many $A$ individuals in the population is simply $i/N$ [79]. Hence, at any time before consensus there is a probability $(N - i)/N$ that 0 will dominate. The expected number of interactions per individual needed to reach consensus is proportional to the population size $N$ [80].

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





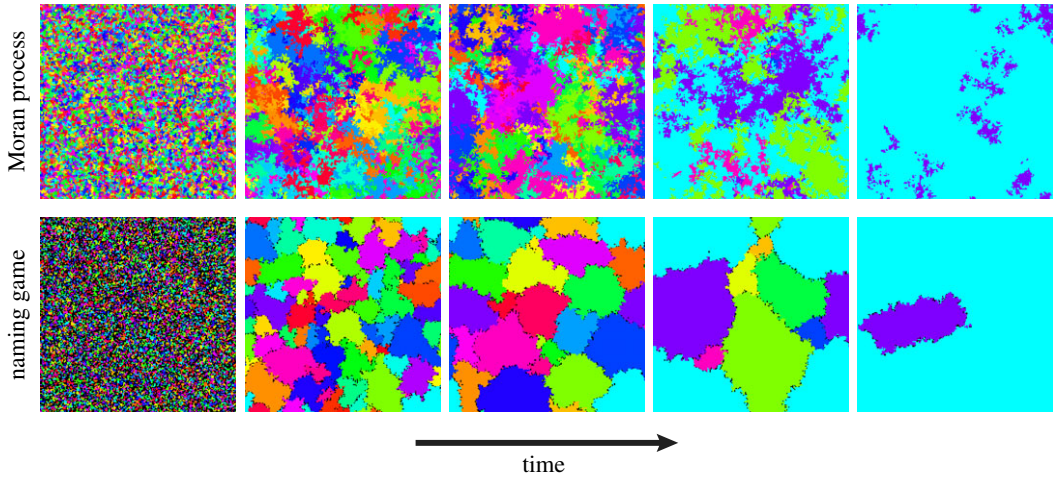


**Figure 2.** From local to global consensus in spatial networks. Snapshots of the temporal evolution of the Moran process (top, simple contagion) and naming game (bottom, complex contagion) on a two-dimensional lattice with coordination number 4 and periodic boundary conditions. While compact clusters of agreeing agents form in the naming game, in the Moran process regions of the same colour are difficult to identify and often broken in more pieces. Population of $N = 40\,000$ agents, initial condition with $M = N$ different states (i.e. each agent starts in a different state). Colours correspond to different states, with the exception of the left panels where for visualization purposes it is possible that different states are rendered in the same colour. Black points in the naming game correspond to agents with more than one name in their inventory.

In the binary (or 'two-state') naming game, agents are initially assigned with one of two names (e.g. $A$ or $B$) and can find themselves in one of the three states identified by an inventory that contains only name $A$, only name $B$ or both $A$ and $B$ [81]. A contact between $A$ ($B$) and $AB$ will increase the population of $A$ ($B$) with probability $p = \frac{3}{4}$. Thus, the larger the fraction $n_A$ ($n_B$) of individuals who only know name $A$ ($B$) the more that fraction will increase. Mathematically, the difference between $n_A$ and $n_B$ (note that $n_{AB} = 1 - n_A - n_B$) evolves according to $\mathrm{d}(n_A - n_B)/\mathrm{d}t \propto (n_A - n_B)$, meaning that the larger faction will always impose its consensus in large populations [81]. When the number of states is not restricted, the dynamics is characterized by an initial phase of competition between names, followed by a winner-take-all regime in which the most popular convention progressively eliminates all the competitors [34], in a process known as *symmetry breaking* (figure 1). The time needed to reach consensus is faster than in the Moran process, and proportional to $\log N$ and $\sqrt{N}$ interactions for the binary and the unrestricted models, respectively [34,82].

## 5.2. Spatial networks

On two-dimensional regular lattices the time required to reach consensus is $t_{\mathrm{consesus}} \sim \ln N$ for the Moran process [80] and $t_{\mathrm{consesus}} \sim N$ for the naming game [83]. While figure 1 might suggest that the dynamics of the two models is similar on lattices, figure 2 shows that important differences exist. In the naming game, local consensus between neighbouring individuals emerges rapidly but different regions reach a consensus on different conventions. Clusters of *local* consensus stay compact and the dynamics proceeds through cluster–cluster competition at the frontier between different regions. In the Moran process, on the other hand, simple contagion prevents the formation of such compact clusters, and the path to global consensus is dominated by fluctuations as in homogeneously mixing populations.

It is important to note that, beyond these two models, other scenarios exist and opposite results can be found. For example, in the context of the coordination game with bounded rationality [56], convergence to the risk-dominant strategy is slow on fully connected graphs, where initial conditions play a predominant role, while evolutionary forces determine the outcome when players interact with small sets of neighbours in clustered networks [84].

## 5.3. Complex networks

Most networks observed in nature are characterized by the small-world property [85,86], describing the fact that the average distance between any pair of nodes grows as the logarithm of the system size,

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





and a broad distribution of node connectivity $k$ [87], often compatible with a scale-free behaviour $P(k) \sim k^{-\gamma}$ with $2 < \gamma < 3$ [87–89]. On such topologies, both the Moran process and the naming game recover the behaviour and scaling exponents observed in homogeneously mixing populations [69,70,90]. In the naming game, after an initial phase of local agreement, the small-world property favours the spreading of conventions between different regions thus preventing the formation of regional clusters [91].

However, on scale-free networks Moran and voter model behave differently, the presence of hubs slowing down consensus in the Moran process and favouring it for the voter model [69–72,92]. In general, in any model describing pairwise interactions, the topology and role of the agents become entangled on heterogeneous networks. The first individual is selected according to the degree distribution $P(k)$, while the second individual, being selected among the neighbours of the first one, is sampled from a different distribution, which in the case of uncorrelated networks is $Q(k) \sim kP(k)$ [93].

Of course, further possibilities exist and predictions of game-theoretic models may be antithetic to the ones described above. For example, innovations spread quickly in locally connected networks and geographical networks, while hubs are an obstacle to the spreading of a risk-dominant strategy in a model where the payoff of each alternative increases with the number of neighbours who are adopting the same choice [94].

## 6. Fragile consensus and committed minorities

The large majority of models describe consensus as an absorbing state: once reached, it will persist indefinitely [8,95]. However, social consensus is often fragile. Apparently, small shocks or weak forces can result in global shifts of behaviour, causing consensus to move from one equilibrium to a different one. Cohabitation of unmarried couples, same-sex relationships and social attitudes towards legal and illegal drugs have changed over the course of the last decades [35,96]. Interestingly, often the transition from one equilibrium to the new one is swift, and the reshaping of consensus can be described in terms of physical concepts such as a phase transition [22,97] or a collective swing due to spontaneous fluctuations [98].

An important question is whether a small fraction of committed actors can push the majority of the population towards a different equilibrium. Various social phenomena, from revolutions [99] to the constant renewal of current day slang [100] and to fashions and fads [35] are in fact attributed to the activity of initially small groups. The two models we have examined above have been extensively studied in this context, following pioneering insights from different approaches [101].

In the case of the Moran process, even a minority of non-committed individuals has always a chance to sway the majority opinion. It is remarkable, however, that a single committed agent is able to lead the whole population towards the state it chooses in spatial lattices, while it is unable to do so in higher dimensions [102,103]. In the binary naming game, on the other hand, we have seen that the majority opinion will always be imposed at the population level. However, it can be shown that a minority of individuals committed on name $B$ will be able to flip the consensus reached on $A$ provided that its size exceeds a threshold of around 10% of the individuals [104–106]. A similar threshold has been observed also in radically different models [107], while a more heterogeneous distribution of individual commitments yields to minority thresholds in the range between 10 and 40% in the context of the naming game [108].

## 7. Obstacles to spontaneous consensus and coexistence of different states

Most formal models of social influence seem to imply that consensus is unavoidable [109–112]. However, disagreement characterizes many aspects of our society. A natural question is therefore what factors can hinder the process of consensus in models that would otherwise lead to it.

One natural answer is topology. Networks characterized by a strong community structure can enormously slow down, or even prevent, consensus in models of complex contagion [90,113,114]. The same mechanisms yielding compact clusters of agreeing individuals in spatial networks guarantee the cohesiveness of a topological community. Furthermore, simple modifications of the microscopic rules, such as a nonlinear dependence of the transition rates on the states of neighbouring nodes in the voter model [115] or an irresolute attitude of the agents in the naming game [81], may guarantee the coexistence of different states even on lattices or fully connected graphs.

A different mechanism is proposed by the well-known Axelrod model of dissemination of culture [116], defined as a set of individual attributes that are subject to social influence. Given that individuals have a tendency to interact more with others who share their opinion (homophily) and that

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





**Table 1.** Outstanding questions.

- Can behavioural change be engineered? Can we foster social consensus on beneficial behavioural norms, such as practices of environmental sustainability or social inclusion? Conversely, how can negative yet widespread norms—from bullying to corruption—be eradicated?
- How can we contrast the formation of online self-organizing communities, or 'echo-chambers'? How can the connectivity of a social network be increased? How robust are these echo-chambers? How do overlapping echo-chambers interact?
- How can committed minorities be put to use to induce social change? Can their role be tested in the laboratory? Are the properties of the network structure in social interactions a key factor for the effectiveness of committed minorities?
- How are online social networks changing the mechanisms of social consensus? What is the interaction between online and offline paths to consensus? What is the role of centralized and decentralized mechanisms of information production on the formation of social consensus? How does consensus emerge on new ways to attribute and store values such as cryptocurrencies?

interactions between individuals tend to increase their similarity (social influence), where do cultural differences come from? The answer has to be sought in the mechanisms of 'bounded confidence' according to which only individuals that are already sufficiently similar interact [117]. In the model, individuals are characterized by $F$ cultural features that can assume $q$ traits. At each time step, two individuals are randomly selected and interact with a probability proportional to the number of features for which they share the same trait. The result of an interaction is that the two individuals will increase their similarity by aligning one feature for which traits are different. If the number of possible traits, $q$, is small the process will end up in a state of consensus where all individuals share the same trait for the same feature, but a threshold value exists such that for $q > q_c$ consensus will not be reached [118,119]. Furthermore, it has been shown that the interplay between local interactions and the homogenizing effect of a centralized ordering effort produces non-trivial results and may increase the disorder of the system [120,121].

Interestingly, it has recently been shown that topology and homophily interact in online social networks, where users have the possibility to control who to connect to (see [122] for the modelling of this feature in the context of the Axelrod model). Here, tightly connected and relatively isolated communities emerge spontaneously, maintaining and promoting group polarization. These 'echo-chambers' hinder consensus not only at the level of social conventions and norms, but also on the recognition of e.g. scientific evidence [123] with consequences on public debate [124]. Theoretical approaches including dynamical network modelling along with homophily and social influence confirm this picture [125]. An open question, whose urgency has been stressed also by the World Economic Forum, is what can be done to favour a less polarized debate in our society (see also table 1) [126,127].

## 8. Empirical studies

Insights on the emergence of consensus often come from studies designed with a different focus. This section covers some examples representative of different approaches, motivations and implementation schemes.

Language is a natural environment for study of the spontaneous emergence of conventions. While the space of alternatives is naturally rich, however, experiments in this context have often aimed to explore coordination on higher linguistic features (e.g. the emergence of compositionality) involving small population sizes. Galantucci [128] investigated the emergence of a communication code in a simple coordination game. Pairs of physically separated individuals had to coordinate on where to go in order to meet in a simple set of communicating rooms they saw on a screen. Communication was mediated by a system that does not allow users to write (a sliding trackpad). The author found that a communication system emerged, signs could originate from different mappings (movement, position, etc.), systems developed parsimoniously (new signs are related to already established signs) and final signs were well distinct. Related yet different experiments showed that an unstable environment may facilitate the emergence of sophisticated forms of coordination, such as a compositional code, when pairs of individuals communicate [129]. Garrod & Doherty [38] analysed the role of a community, as opposed to just two communicating individuals, where individuals—interacting in pairs—had to describe their changing position in a maze. The presence of more users (up to $N = 10$, in the experiments) slowed down the initial agreement, but resulted in a more stable consensus, i.e. in a final state with more successful

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





interactions based on more stable codes, in agreement with the Lewis view of conventions as solutions to *collective* coordination problems [1,130,131].

The spontaneous emergence of consensus was explicitly addressed in Centola & Baronchelli [132] through a coordination game played by group sizes of up to $N = 96$ individuals. In a given round of the game, two network 'neighbours' were chosen at random to play with one another. Both players simultaneously assigned names to a human face. If the players coordinated on a name, they were rewarded with a successful payment, otherwise they were penalized. After a single round, the participants could see only the choices that they and their partner had made. They were then randomly assigned to play with a new neighbour in their social network, and a new round would begin. The object (i.e. the human face) that participants were trying to name was the same for the entire duration of the game, and for all members of the game. The experiments showed that global consensus emerges in homogeneously mixing populations, while different clusters of local consensus appear in spatial networks in agreement with the predictions of the naming game model [34].

Kearns *et al.* [25] explored the problem of consensus in the presence of an explicit incentive for collective agreement. A population of $N = 36$ individuals was arranged on networks with different topologies to play a networked version of the classic 'Battle of the Sexes' game. Individuals were in one of two possible states, labelled 'red' and 'blue', and their payoff in the game depended on which state will eventually be adopted by the whole population. Each individual knew the state of their neighbours in the network and could change colour at each time step. Results showed that when incentives were randomly distributed in the population, so that 50% of the individuals prefer red blue and the other 50% prefer red red, consensus was reached in only 57% of trials. When, on the other hand, a certain payoff was assigned to a minority of individuals occupying well-connected nodes in a heterogeneous network consensus is much more likely (89% of trials). Judd *et al.* [133] adopted a similar setting, in experiments where global agreement was the explicit goal and individuals characterized by a simple colour variable have access to the state of their neighbours. Starting from a network characterized by a strong community structure (six communities for a population of $N = 36$ individuals), these experiments confirmed that 'long-distance' connections, i.e. the small-world property, promote consensus.

Empirical investigations of the spreading of behaviour have provided important insights on the existence and nature of complex contagions, which as we have seen is a crucial ingredient of many models for the emergence of consensus. Microscopic complex contagion has been studied in the laboratory [134] and in offline [135] and online [136,137] social networks also in relation with its interplay with the topology of the network. Other experiments have started to unveil previously neglected aspects of the coordination process. For example, the structure of incentives has been investigated, showing that higher stakes increase the pressure to establish and adhere to shared expectations that persist across rounds [138].

Finally, outside of the laboratory, conventions have been investigated for example using Twitter. Focusing on the adoption (i.e. first use) of markers for retweet or tweet quoting, Kooti *et al*. [139] found that, despite many alternatives being proposed eventually the conventions of 'RT' and 'via' became dominant. Interestingly, successful conventions were initially proposed and adopted by active and well-connected users at the core of the Twitter community, showing that status, influence and connectedness play an important role, changing the ideal condition of interacting *peers* [139]. Interestingly, a similar role of earliest users in determining the normative consensus has been found also in Wikipedia [140].

# 9. Concluding remarks and outlook

This overview has necessarily been (very) selective, but it allows us to draw encouraging conclusions. Major advancements occurring in the past years have shed new light on the process of consensus formation. Theoretical milestones in game theory and complexity science have benefited by the steady increase of computational power and the consequent investigation of a large number of models for the study of consensus formation. Different hypotheses have been tested and the micro–macro connection is now much clearer in many situations, although important questions remain open (table 1). Very recently, finally, empirical approaches, the analysis of human activity on social media and the use of wearable sensors have started to shed light on the mechanisms at play in our society. It is likely that further insights will be produced by the synergy of these three approaches in the next few years.

Data accessibility. As a review article, there are no additional data. Data referred to can be found in the appropriate cited reference.

Competing interests. I declare I have no competing interests.

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018

Funding. We received no funding for this study.
Acknowledgements. The author is grateful to Prof. Frank Schweitzer for very helpful comments on the manuscript.





# Appendix A

This section provides a short definition of some of the terms used in the main text. In some cases, a broader definition exists but only the one useful to an easier reading of the present paper is provided.

*Basin of attraction*: A region of the phase space of a dynamical system such that initial conditions chosen in that region dynamically evolve to a particular attractor.
*Contagion*: Transmission of a disease, idea or behaviour from a person to another by close contact.

- *Simple contagion*: Process in which successive exposures to a pathogen or behaviour are independent and characterized by the same probability $p$ of infection.
- *Complex contagion*: Process in which the probability of infection (i.e. for example, adoption of a behaviour) depends on the number of exposures in a complex, nonlinear, way.

*Nash equilibrium*: Stable state of a system of interacting individuals in which no player can benefit by changing strategies while the other players keep theirs unchanged.
*Self-organization*: The capability of a system to acquire a functional, spatial or temporal structure without specific interference from the outside [141]. Sometimes identified with 'spontaneous' order in the Social Sciences.
*(Spontaneous) Symmetry breaking*: Process of symmetry reduction in a system evolving according to symmetric laws. Arbitrarily small fluctuations drive the system out of the initially symmetrical state and into a final asymmetrical state.

Topology-related terms:

*Homogeneously mixing population*: Population in which agents occupy the vertices of a complete, or fully connected, graph.
*Lattice*: An arrangement in space of isolated points (lattice points) in a regular pattern. In two dimensions, the word 'lattice' is typically used to refer to a regular grid in which each point is connected to four neighbouring points.
*Network*: A collection of points, called nodes, joined by lines, referred to as links. Vertices represent the elementary components of a system, for example the individuals in a population, whereas links stand for the possible interactions between pairs of components (see also [89,93,142,143] for more details on the quantities detailed below).

- *Community*: Although many definitions exist, a community can be generally defined as a set of nodes which are more tightly connected with one another than with other nodes in the network [144].
- *Degree of a node*: The degree $k_i$ of a node $i$ is defined as the number of other nodes to which it is connected, i.e. to the number of its 'neighbours'.
- *Degree distribution*: The probability $P(k)$ that a randomly chosen vertex has degree $k$.
- *Heterogeneous, or 'scale free', networks*: Networks with a heavy-tailed degree distribution that can often be approximated by a power-law, $P(k) \sim k^{-\gamma}$, with $\gamma$ typically between 2 and 3. The presence of extremely well-connected nodes, or 'hubs', is responsible for many of the interesting properties of real-world networks.
- *Homogeneous networks*: Networks with a well-peaked and exponentially decaying degree-distribution, where the variation in connectivity among nodes is limited and hubs are absent.
- *Shortest path length*, or distance, between vertices $i$ and $j$ is the length (in number of edges) of the shortest path joining $i$ and $j$.
- *Small-world property*: A property shown by many real networks that exhibit a small value of the average shortest path length, increasing with network size logarithmically or slower. This property is in stark contrast to the larger diameter of regular lattices, which grows algebraically with lattice size.

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018





# References


1. Lewis D. 1969 *Convention: a philosophical study*. Oxford, UK: Blackwell.
2. Hechter M, Horne C. 2003 *Theories of social order: a reader*. Palo Alto, CA: Stanford University Press.
3. Attanasi A *et al.* 2014 Information transfer and behavioural inertia in starling flocks. *Nat. Phys.* **10**, 691–696. (doi:10.1038/nphys3035)
4. Rosenthal SB, Twomey CR, Hartnett AT, Wu HS, Couzin ID. 2015 Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proc. Natl Acad. Sci. USA* **112**, 4690–4695. (doi:10.1073/pnas.1420068112)
5. Peierls R. 1936 On Ising's model of ferromagnetism. *Mathematical Proc. Cambridge Phil. Soc.* **32**, 477–481.
6. Werfel J, Petersen K, Nagpal R. 2014 Designing collective behavior in a termite-inspired robot construction team. *Science* **343**, 754–758. (doi:10.1126/science.1245842)
7. Skyrms B. 1996 *Evolution of the social contract*. Cambridge, UK: Cambridge University Press.
8. Castellano C, Fortunato S, Loreto V. 2009 Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591–646. (doi:10.1103/RevModPhys.81.591)
9. Ehrlich PR, Levin SA. 2005 The evolution of norms. *PLoS Biol.* **3**, e194. (doi:10.1371/journal.pbio.0030194)
10. Young HP. 2015 The evolution of social norms. *Annu. Rev. Econ.* **7**, 359–387. (doi:10.1146/annurev-economics-080614-115322)
11. Nyborg K *et al.* 2016 Social norms as solutions. *Science* **354**, 42–43. (doi:10.1126/science.aaf8317)
12. Collins English Dictionary. 2017 See Dictionary.com (accessed 15 April 2017).
13. Sedley D. 2003 *Plato's Cratylus*. Cambridge, UK: Cambridge University Press.
14. Jacob B, Jacob E, Jacob W. 2007 *The first book of the Bible, Genesis*. New York, NY: KTAV Publishing House Inc.
15. Young HP. 1996 The economics of convention. *J. Econ. Perspect.* **10**, 105–122. (doi:10.1257/jep.10.2.105)
16. Durkheim E. 1976 *The elementary forms of religious life*. London, UK: Routledge.
17. Mead GH. 1934 *Mind, self and society*, vol. 111. Chicago, IL: University of Chicago Press.
18. O'Malley J, Davis. RA. 1994 *Marx: early political writings*. Cambridge, UK: Cambridge University Press.
19. Pinker S. 1995 *The language instinct: the new science of language and mind*, vol. 7529. Harmondsworth, UK: Penguin.
20. Hobbes T. 2006 *Leviathan*. London, UK: A&C Black.
21. Myers JH, Robertson TS. 1972 Dimensions of opinion leadership. *J. Mark. Res.* **9**, 41–46. (doi:10.2307/3149604)
22. Gladwell M. 2000 *The tipping point: how little things can make a big difference*. Boston, MA: Little Brown.
23. McCombs ME, Shaw DL. 1972 The agenda-setting function of mass media. *Public. Opin. Q.* **36**, 176–187. (doi:10.1086/267990)
24. McCombs M. 1997 Building consensus: The news media's agenda-setting roles. *Political Commun.* **14**, 433–443. (doi:10.1080/105846097199236)
25. Kearns M, Judd S, Tan J, Wortman J. 2009 Behavioral experiments on biased voting in networks. *Proc. Natl Acad. Sci. USA* **106**, 1347–1352. (doi:10.1073/pnas.0808147106)
26. Oliver PE. 1993 Formal models of collective action. *Annu. Rev. Sociol.* **19**, 271–300. (doi:10.1146/annurev.so.19.080193.001415)
27. Merton. RK. 1968 *Social theory and social structure: toward the codification of theory and research*. London, UK: Simon and Schuster.
28. Asch SE. 1955 Opinions and social pressure. *Read. Soc. Anim.* **193**, 17–26. (doi:10.1038/scientificamerican1155-31)
29. Sugden R. 1989 Spontaneous order. *J. Econ. Perspect.* **3**, 85–97. (doi:10.1257/jep.3.4.85)
30. Axelrod R. 1986 An evolutionary approach to norms. *Am. Political Sci. Rev.* **80**, 1095–1111. (doi:10.2307/1960858)
31. Young HP. 1993 The evolution of conventions. *Econometrica: J. Econ. Soc.* **61**, 57–84.
32. Steels L. 1995 A self-organizing spatial vocabulary. *Artif. Life.* **2**, 319–332. (doi:10.1162/artl.1995.2.319)
33. Bicchieri C. 2005 *The grammar of society: the nature and dynamics of social norms*. Cambridge, UK: Cambridge University Press.
34. Baronchelli A, Felici M, Loreto V, Caglioti E, Steels L. 2006 Sharp transition towards shared vocabularies in multi-agent systems. *J. Stat. Mech.: Theory Exp.* **2006**, P06014. (doi:10.1088/1742-5468/2006/06/P06014)
35. Bikhchandani S, Hirshleifer D, Welch I. 1992 A theory of fads, fashion, custom, and cultural change as informational cascades. *J. Political Econ.* **100**, 992–1026. (doi:10.1086/261849)
36. Rogers EM. 2010 *Diffusion of innovations*. New York, NY: Simon and Schuster.
37. Wittgenstein L. 1958 *Philosophical investigations*. Translation by GEM Anscombe. Oxford, UK: Blackwell.
38. Garrod S, Doherty G. 1994 Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistic conventions. *Cognition* **53**, 181–215. (doi:10.1016/0010-0277(94)90048-5)
39. Boyd R, Richerson PJ. 1992 Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethol. Sociobiol.* **13**, 171–195. (doi:10.1016/0162-3095(92)90032-Y)
40. Schelling TC. 1960 *The strategy of conflict*. Cambridge, MA: Harvard University Press.
41. Dybvig PH, Spatt CS. 1983 Adoption externalities as public goods. *J. Public. Econ.* **20**, 231–247. (doi:10.1016/0047-2727(83)90012-9)
42. Farrell J, Saloner G. 1985 Standardization, compatibility, and innovation. *Rand. J. Econ.* **16**, 70–83. (doi:10.2307/2555589)
43. Katz ML, Shapiro C. 1985 Network externalities, competition, and compatibility. *Am. Econ. Rev.* **75**, 424–440.
44. Arthur WB. 1989 Competing technologies, increasing returns, and lock-in by historical events. *Econ. J.* **99**, 116–131. (doi:10.2307/2234208)
45. Valverde S, Solé RV. 2015 Punctuated equilibrium in the large-scale evolution of programming languages. *J. R. Soc. Interface* **12**, 20150249. (doi:10.1098/rsif.2015.0249)
46. Jones SRG. 1984 *The economics of conformism*. Oxford, UK: Blackwell.
47. Richerson PJ, Boyd R. 2005 *Not by genes alone: how culture transformed human evolution*. Chicago, IL: Chicago University Press.
48. Whiten A, Horner V, De Waal FBM. 2005 Conformity to cultural norms of tool use in chimpanzees. *Nature* **437**, 737–740. (doi:10.1038/nature04047)
49. Miller MB, Bassler BL. 2001 Quorum sensing in bacteria. *Annu. Rev. Microbiol.* **55**, 165–199. (doi:10.1146/annurev.micro.55.1.165)
50. Mallon E, Pratt S, Franks N. 2001 Individual and collective decision-making during nest site selection by the ant *Leptothorax albipennis*. *Behav. Ecol. Sociobiol.* **50**, 352–359. (doi:10.1007/s002650100377)
51. Seeley T, Visscher PK. 2004 Group decision making in nest-site selection by honey bees. *Apidologie* **35**, 101–116. (doi:10.1051/apido:2004004)
52. Harsanyi JC, Selten R. 1988 *A general theory of equilibrium selection in games*, vol. 1. Cambridge, MA: MIT Press Books.
53. Von Neumann J, Morgenstern O. 1947 *Theory of games and economic behavior*. 2nd edn. Princeton, NJ: Princeton University Press.
54. Ellison G. 2000 Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution. *Rev. Econ. Stud.* **67**, 17–45. (doi:10.1111/1467-937X.00119)
55. Samuelson L. 1998 *Evolutionary games and equilibrium selection*, vol. 1. Cambridge, MA: MIT Press.
56. Kandori M, Mailath GJ, Rob R. 1993 Learning, mutation, and long run equilibria in games. *Econometrica: J. Econ. Soc.* **61**, 29–56. (doi:10.2307/2951777)
57. Binmore K. 1987 Modeling rational players: Part I. *Econ. Philos.* **3**, 179–214. (doi:10.1017/S0266267100002893)
58. Simon HA. 1996 *The sciences of the artificial*, 3rd edn. Cambridge, MA: MIT Press.
59. Maynard Smith J, Price GR. 1973 The logic of animal conflict. *Nature* **246**, 15–18. (doi:10.1038/246015a0)
60. Foster D, Young P. 1990 Stochastic evolutionary game dynamics? *Theor. Popul. Biol.* **38**, 219–232.
61. Schelling TC. 1971 Dynamic models of segregation. *J. Math. Sociol.* **1**, 143–186. (doi:10.1080/0022250X.1971.9989794)
62. Axelrod RM. 1997 *The complexity of cooperation: agent-based models of competition and collaboration*. Princeton, NJ: Princeton University Press.
63. Reynolds CW. 1987 Flocks, herds and schools: a distributed behavioral model. *ACM SIGGRAPH Comput. Graph.* **21**, 25–34. (doi:10.1145/37402.37406)
64. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. 2015 Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925. (doi:10.1103/RevModPhys.87.925)

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018






65. Centola D, Macy M. 2007 Complex contagions and the weakness of long ties. *Am. J. Sociol.* **113**, 702–734. (doi:10.1086/521848)
66. Moran PAP. 1958 2007 Random processes in genetics. *Mathematical Proc. Cambridge Phil. Soc.* **54**, 60–71.
67. Clifford P, Sudbury A. 1973 A model for spatial conflict. *Biometrika* **60**, 581–588. (doi:10.1093/biomet/60.3.581)
68. Holley RA, Liggett TM. 1975 Ergodic theorems for weakly interacting infinite systems and the voter model. *Ann. Probab.* **3**, 643–663. (doi:10.1214/aop/1176996306)
69. Suchecki K, Eguíluz VM, San Miguel M. 2005 Voter model dynamics in complex networks: role of dimensionality, disorder, and degree distribution. *Phys. Rev. E* **72**, 036132. (doi:10.1103/PhysRevE.72.036132)
70. Castellano C. 2005 Effect of network topology on the ordering dynamics of voter models. In *AIP Conf. Proceedings*, vol. 779, pp. 114–120. American Institute of Physics.
71. Sood V, Redner S. 2005 Voter model on heterogeneous graphs. *Phys. Rev. Lett.* **94**, 178701. (doi:10.1103/PhysRevLett.94.178701)
72. Sood V, Antal T, Redner S. 2008 Voter models on heterogeneous networks. *Phys. Rev. E* **77**, 041121. (doi:10.1103/PhysRevE.77.041121)
73. Barr DJ. 2004 Establishing conventional communication systems: is common knowledge necessary? *Cogn. Sci.* **28**, 937–962. (doi:10.1016/j.cogsci.2004.07.002)
74. Latane B. 1981 The psychology of social impact. *Am. Psychol.* **36**, 343. (doi:10.1037/0003-066X.36.4.343)
75. Nowak A, Szamrej J, Latané B. 1990 From private attitude to public opinion: a dynamic theory of social impact. *Psychol. Rev.* **97**, 362–376. (doi:10.1037/0033-295X.97.3.362)
76. Lewenstein M, Nowak A, Latané B. 1992 Statistical mechanics of social impact. *Phys. Rev. A* **45**, 763. (doi:10.1103/PhysRevA.45.763)
77. Hołyst JA, Kacperski K, Schweitzer F. 2000 Phase transitions in social impact models of opinion formation. *Phys. A: Stat. Mech. Appl.* **285**, 199–210. (doi:10.1016/S0378-4371(00)00282-X)
78. Schweitzer F, Hołyst JA. 2000 Modelling collective opinion formation by means of active brownian particles. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **15**, 723–732. (doi:10.1007/s100510051177)
79. Nowak MA. 2006 *Evolutionary dynamics*. Cambridge, MA: Harvard University Press.
80. Cox JT. 1989 Coalescing random walks and voter model consensus times on the torus in Zd. *Ann. Probab.* **17**, 1333–1366.
81. Baronchelli A, Dall'Asta L, Barrat A, Loreto V. 2007 Nonequilibrium phase transition in negotiation dynamics. *Phys. Rev. E* **76**, 051102. (doi:10.1103/PhysRevE.76.051102)
82. Baronchelli A, Loreto V, Steels L. 2008 In-depth analysis of the naming game dynamics: the homogeneous mixing case. *Int. J. Mod. Phys. C* **19**, 785–812. (doi:10.1142/S0129183108012522)
83. Baronchelli A, Dall'Asta L, Barrat A, Loreto V. 2006 Topology-induced coarsening in language games. *Phys. Rev. E* **73**, 015102. (doi:10.1103/PhysRevE.73.015102)
84. Ellison G. 1993 Learning, local interaction, and coordination. *Econometrica: J. Econ. Soc.* **61**, 1047–1071. (doi:10.2307/2951493)
85. Travers J, Milgram S. 1967 The small world problem. *Phychol. Today* **1**, 61–67.
86. Watts DJ, Strogatz SH. 1998 Collective dynamics of 'small-world' networks. *Nature* **393**, 440–442. (doi:10.1038/30918)
87. Barabási A-L, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512. (doi:10.1126/science.286.5439.509)
88. Newman MEJ. 2003 The structure and function of complex networks. *SIAM Rev.* **45**, 167–256. (doi:10.1137/S003614450342480)
89. Caldarelli G. 2007 *Scale-free networks: complex webs in nature and technology*. Oxford, UK: Oxford University Press.
90. Dall'Asta L, Baronchelli A, Barrat A, Loreto V. 2006 Nonequilibrium dynamics of language games on complex networks. *Phys. Rev. E* **74**, 036105. (doi:10.1103/PhysRevE.74.036105)
91. Dall'Asta L, Baronchelli A, Barrat A, Loreto V. 2006 Agreement dynamics on small-world networks. *EPL* (*Europhys. Lett.*) **73**, 969. (doi:10.1209/epl/i2005-10481-7)
92. Vazquez F, Eguíluz VM. 2008 Analytical solution of the voter model on uncorrelated networks. *New. J. Phys.* **10**, 063011. (doi:10.1088/1367-2630/10/6/063011)
93. Newman M. 2010 *Networks: an introduction*. New York. Oxford University Press Inc.
94. Montanari A, Saberi A. 2010 The spread of innovations in social networks. *Proc. Natl Acad. Sci. USA* **107**, 20 196–20 201. (doi:10.1073/pnas.1004098107)
95. Galam S. 2008 Sociophysics: a review of galam models. *Int. J. Modern Phys. C* **19**, 409–440. (doi:10.1142/S0129183108012297)
96. Kuran T. 1989 Sparks and prairie fires: a theory of unanticipated political revolution. *Public Choice* **61**, 41–74. (doi:10.1007/BF00116762)
97. Haken H. 1975 Cooperative phenomena in systems far from thermal equilibrium and in nonphysical systems. *Rev. Mod. Phys.* **47**, 67–121. (doi:10.1103/RevModPhys.47.67)
98. Cavagna A, Giardina I, Jelic A, Silvestri E, Viale M. 2017 Non-symmetric interactions trigger collective swings in globally ordered systems. *Phys. Rev. Lett.* **118**, 138003. (doi:10.1103/PhysRevLett.118.138003)
99. Gladwell M. 2010 Small change. *The New Yorker* **4**, 42–49.
100. Lighter JE, House R. 1994 *Random House historical dictionary of American slang*. New York, NY: Random House.
101. Galam S, Jacobs F. 2007 The role of inflexible minorities in the breaking of democratic opinion dynamics. *Phys. A: Stat. Mech. Appl.* **381**, 366–376. (doi:10.1016/j.physa.2007.03.034)
102. Mobilia M. 2003 Does a single zealot affect an infinite group of voters? *Phys. Rev. Lett.* **91**, 028701. (doi:10.1103/PhysRevLett.91.028701)
103. Mobilia M, Redner S. 2003 Majority versus minority dynamics: phase transition in an interacting two-state spin system. *Phys. Rev. E* **68**, 046106. (doi:10.1103/PhysRevE.68.046106)
104. Xie J, Sreenivasan S, Korniss G, Zhang W, Lim C, Szymanski BK. 2011 Social consensus through the influence of committed minorities. *Phys. Rev. E* **84**, 011130. (doi:10.1103/PhysRevE.84.011130)
105. Marvel SA, Hong H, Papush A, Strogatz SH. 2012 Encouraging moderation: clues from a simple model of ideological conflict. *Phys. Rev. Lett.* **109**, 118702. (doi:10.1103/PhysRevLett.109.118702)
106. Mistry D, Zhang Q, Perra N, Baronchelli A. 2015 Committed activists and the reshaping of status-quo social consensus. *Phys. Rev. E* **92**, 042805. (doi:10.1103/PhysRevE.92.042805)
107. Halu A, Zhao K, Baronchelli A, Bianconi G. 2013 Connect and win: the role of social networks in political elections. *EPL* (*Europhys. Lett.*) **102**, 16002. (doi:10.1209/0295-5075/102/16002)
108. Niu X, Doyle C, Korniss G, Szymanski BK. 2017 The impact of variable commitment in the naming game on consensus formation. *Sci. Rep.* **7**, 41750. (doi:10.1038/srep41750)
109. Shao J, Havlin S, Stanley HE. 2009 Dynamic opinion model and invasion percolation. *Phys. Rev. Lett.* **103**, 018701. (doi:10.1103/PhysRevLett.103.018701)
110. Mäs M, Flache A, Helbing D. 2010 Individualization as driving force of clustering phenomena in humans. *PLoS Comput. Biol.* **6**, e1000959. (doi:10.1371/journal.pcbi.1000959)
111. Abelson RP. 1964 Mathematical models of the distribution of attitudes under controversy. *Contrib. Math. Psychol.* **14**, 1–160.
112. Li Q, Braunstein LA, Wang H, Shao J, Stanley HE, Havlin S. 2013 Non-consensus opinion models on complex networks. *J. Stat. Phys.* **151**, 92–112. (doi:10.1007/s10955-012-0625-4)
113. Lambiotte R, Ausloos M. 2007 Coexistence of opposite opinions in a network with communities. *J. Stat. Mech.: Theory Exp.* **2007**, P08026. (doi:10.1088/1742-5468/2007/08/P08026)
114. Lu Q, Korniss G, Szymanski BK. 2009 The naming game in social networks: community formation and consensus engineering. *J. Econ. Interact. Coord.* **4**, 221–235. (doi:10.1007/s11403-009-0057-7)
115. Schweitzer F, Behera L. 2009 Nonlinear voter models: the transition from invasion to coexistence. *Eur. Phys. J. B-Condens. Matter Complex Syst.* **67**, 301–318. (doi:10.1140/epjb/e2009-00001-3)
116. Axelrod R. 1997 The dissemination of culture a model with local convergence and global polarization. *J. Conflict Resolution* **41**, 203–226. (doi:10.1177/0022002797041002001)
117. Deffuant G, Neau D, Amblard F, Weisbuch G. 2000 Mixing beliefs among interacting agents. *Adv. Complex Syst.* **3**, 87–98. (doi:10.1142/S0219525900000078)
118. Castellano C, Marsili M, Vespignani A. 2000 Nonequilibrium phase transition in a model for social influence. *Phys. Rev. Lett.* **85**, 3536. (doi:10.1103/PhysRevLett.85.3536)
119. Klemm K, Eguíluz VM, Toral R, San Miguel M. 2003 Nonequilibrium transitions in complex networks: a model of social interaction. *Phys. Rev. E* **67**, 026120. (doi:10.1103/PhysRevE.67.026120)
120. González-Avella JC, Cosenza MG, Tucci K. 2005 Nonequilibrium transition induced by mass media in a model for social influence. *Phys. Rev. E* **72**, 065102. (doi:10.1103/PhysRevE.72.065102)
121. González-Avella JC, Eguíluz VM, Cosenza MG, Klemm K, Herrera JL, San Miguel M. 2006 Local versus global interactions in nonequilibrium

Downloaded from http://rsos.royalsocietypublishing.org/ on June 26, 2018






transitions: a model of social dynamics. *Phys. Rev. E* **73**, 046119. (doi:10.1103/PhysRevE.73.046119)

122. Centola D, Gonzalez-Avella JC, Eguiluz VM, San Miguel M. 2007 Homophily, cultural drift, and the co-evolution of cultural groups. *J. Conflict Resolution* **51**, 905–929. (doi:10.1177/0022002707307632)
123. Del Vicario M, Vivaldo G, Bessi A, Zollo F, Scala A, Caldarelli G, Quattrociocchi W. 2016 Echo chambers: emotional contagion and group polarization on Facebook. *Sci. Rep.* **6**, 37825. (doi:10.1038/srep37825)
124. Del Vicario M, Bessi A, Zollo F, Petroni F, Scala A, Caldarelli G, Stanley HE, Quattrociocchi W. 2016 The spreading of misinformation online. *Proc. Natl Acad. Sci. USA* **113**, 554–559. (doi:10.1073/pnas.1517441113)
125. Starnini M, Frasca M, Baronchelli A. 2016 Emergence of metapopulations and echo chambers in mobile agents. *Sci. Rep.* **6**, 31834. (doi:10.1038/srep31834)
126. Howell L. 2013 Digital wildfires in a hyperconnected world. In *Global Risks 2013*. Geneva, Switzerland: World Economic Forum.
127. Zollo F, Bessi A, Del Vicario M, Scala A, Caldarelli G, Shekhtman L, Havlin S, Quattrociocchi W. 2017 Debunking in a world of tribes. *PLoS ONE* **12**, e0181821. (doi:10.1371/journal.pone.0181821)
128. Galantucci B. 2005 An experimental study of the emergence of human communication systems. *Cogn. Sci.* **29**, 737–767. (doi:10.1207/s15516709cog0000_34)
129. Selten R, Warglien M. 2007 The emergence of simple languages in an experimental coordination game. *Proc. Natl Acad. Sci. USA* **104**, 7361–7366. (doi:10.1073/pnas.0702077104)
130. Lewis D. 1975 Languages and language. In *Minnesota studies in the Philosophy of Science*, vol. 7, pp. 3–35. Minneapolis, MN: University of Minnesota Press.
131. Garrod S, Anderson A. 1987 Saying what you mean in dialogue: a study in conceptual and semantic co-ordination. *Cognition* **27**, 181–218. (doi:10.1016/0010-0277(87)90018-7)
132. Centola D, Baronchelli A. 2015 The spontaneous emergence of conventions: an experimental study of cultural evolution. *Proc. Natl Acad. Sci. USA* **112**, 1989–1994. (doi:10.1073/pnas.1418838112)
133. Judd S, Kearns M, Vorobeychik Y. 2010 Behavioral dynamics and influence in networked coloring and consensus. *Proc. Natl Acad. Sci. USA* **107**, 14 978–14 982. (doi:10.1073/pnas.1001280107)
134. Centola D. 2010 The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197. (doi:10.1126/science.1185231)
135. Christakis NA, Fowler JH. 2007 The spread of obesity in a large social network over 32 years. *N. Engl. J. Med.* **357**, 370–379. (doi:10.1056/NEJMsa066082)
136. Hodas NO, Lerman K. 2014 The simple rules of social contagion. *Sci. Rep.* **4**, 4343. (doi:10.1038/srep04343)
137. Aral S, Nicolaides C. 2017 Exercise contagion in a global social network. *Nat. Commun.* **8**, 14753. (doi:10.1038/ncomms147538)
138. Hawkins RXD, Goldstone RL. 2016 The formation of social conventions in real-time environments. *PLoS. ONE* **11**, e0151670. (doi:10.1371/journal.pone.0151670)
139. Kooti F, Yang H, Cha M, Gummadi PK, Mason WA. 2012 The emergence of conventions in online social networks. In *International AAAI Conference on Web and Social Media* (*ICWSM*), *North America, May 2012*. See https://www.aaai.org/ocs/index.php/ICWSM/ICWSM12/paper/view/4661 (accessed 5 February 2018).
140. Heaberlin B, DeDeo S. 2016 The evolution of Wikipedia's norm network. *Future Internet* **8**, 14. (doi:10.3390/fi8020014)
141. Haken H. 1988 *Information and self-organisation: a macroscopic approach to complex systems*. Berlin, Germany: Springer-Verlag.
142. Albert R, Barabási A-L. 2002 Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47. (doi:10.1103/RevModPhys.74.47)
143. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang D-U. 2006 Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308. (doi:10.1016/j.physrep.2005.10.009)
144. Fortunato S. 2010 Community detection in graphs. *Phys. Rep.* **486**, 75–174. (doi:10.1016/j.physrep.2009.11.002)