



Minerva Access is the Institutional Repository of The University of Melbourne

**Author/s:**

Kashima, Y;Kirley, M;Stivala, A;Robins, G

**Title:**

Modeling cultural dynamics

**Date:**

2017

**Citation:**

Kashima, Y., Kirley, M., Stivala, A. & Robins, G. (2017). Modeling cultural dynamics. Vallacher, RR (Ed.). Read, SJ (Ed.). Nowak, A (Ed.). Computational Social Psychology, (1), pp.281-307. Routledge.

**Persistent Link:**

<https://hdl.handle.net/11343/252754>

## **Modelling cultural dynamics**

Yoshihisa Kashima  
Michael Kirley  
Alexander Stivala  
Garry Robins

### Affiliation

Yoshihisa Kashima  
Melbourne School of Psychological Sciences  
The University of Melbourne  
Parkville, Victoria 3010  
Australia  
Email: [ykashima@unimelb.edu.au](mailto:ykashima@unimelb.edu.au)

Michael Kirley  
Department of Computing and Information Systems,  
The University of Melbourne  
Parkville, Victoria 3010  
Australia  
Email: [mkirley@unimelb.edu.au](mailto:mkirley@unimelb.edu.au)

Alexander Stivala  
Melbourne School of Psychological Sciences  
The University of Melbourne  
Parkville, Victoria 3010  
Australia  
Email: [stivalaa@unimelb.edu.au](mailto:stivalaa@unimelb.edu.au)

Garry Robins  
Melbourne School of Psychological Sciences  
The University of Melbourne  
Parkville, Victoria 3010  
Australia  
Email: [garrylr@unimelb.edu.au](mailto:garrylr@unimelb.edu.au)

### **Abstract**

There is a growing interest in the dynamics of the formation, maintenance, and transformation of culture in different areas of inquiry (e.g., anthropology, biology, computer science, physics, sociology) and policy relevant arenas (e.g., sustainability, intergroup conflict). Despite its traditional interest, a social psychology of cultural dynamics is yet to realize its potential. Included in its methodological toolbox are formal analytical and agent-based models of the movements of culture. In this chapter, we give a broad overview of the emerging field of computational approaches to cultural dynamics and provide a selective review of two of the most prominent research traditions: Axelrod's model of cultural dissemination and the evolution of cooperation in human populations. We conclude with a discussion of the promises and challenges of modelling cultural dynamics.

Culture is an enigma in contemporary social psychology. Despite its brief prominence in the 60s (e.g., Triandis, 1964), social psychological curiosity about culture waned over the heyday of social cognition in the 70s and 80s, only to regain its prominence in the 90s with the publication of Triandis (1989) and Markus and Kitayama (1991) in *Psychological Review*. Over the last two decades and a half, cultural comparative approaches to culture and psychology have accumulated a large body of literature documenting cultural differences around the globe, though primarily comparing East Asian and Western European based cultural groups, and have brought about a number of insights into cultural diversity across humanity. However, this literature has been largely silent on the question about the dynamics of culture, namely, the stability and change of culture, and the mechanisms that drive the trajectories of cultural change over time (see e.g., Kashima & Gelfand, 2012). *Cultural dynamics* is concerned with these questions – how do individuals' thoughts, feelings, and actions in interaction with others in particular contexts generate the movements of culture – its formation, maintenance, and transformation over time (e.g., Kashima, 2000).

At the heart of this lies the fundamental question for social psychology. How do human individuals interact with the actual and imagined others? In so doing, how do we influence each other? Ever since Sherif and Asch, social influence has been a core concern of social psychology; culture is at one level “what social influence influences” (Axelrod, 1997, p. 207). From the concept of human rights to the landing on the moon, and from genocides in intractable intergroup conflicts to anthropogenic climate change, culture enables the astonishing human adaptiveness and achievements in society and in nature, while at times exhibiting surprising maladaptiveness with pathological and tragic consequences in both. With all its glories and failings, the dynamic complexity of human sociality is fundamental to

human culture. Culture enables human sociality; human sociality constitutes culture. Given their interdependency, culture must be an integral part of human social psychology.

Computational modelling – especially agent-based modelling (e.g., Railsback & Grimm, 2012) – is a useful methodological tool in research on cultural dynamics for the development of a coherent theory, derivation of testable hypotheses, understanding of the past and present, as well as forecasting of future possibilities. Because of the number and heterogeneity of actors, the complexity of social interactions, and different timescales involved (social interaction *in situ* may change in a short timespan, but institutions remain stable over a longer period of time), some formal representations (e.g., difference and differential equations) are often useful in theorizing and describing cultural dynamics. However, these very characteristics often make the type of formal analytical methods necessary for the modeling of cultural dynamics outside the training of social psychologists, and difficult, if not impossible, to use (e.g., A. Nowak & Vallacher, 1998). Agent-based modelling approaches, however, enable social psychologists to explicate their assumptions about culture, construct explicit models of cultural dynamics, and explore their implications in a principled fashion. This chapter is designed to provide a broad and selective introduction to diverse literatures on computational approaches to cultural dynamics.

### **Culture and its Dynamics**

Culture is an essentially contested concept – depending on one's theoretical perspective, it can be defined in a multitude of ways. In the current approach, we define culture as *a set of non-genetically transmissible information that is commonly available, accessible, and applicable in a human group*. Cultural information typically takes the form of ideas (e.g., liberty, equality, and fraternity) or practices (e.g., how to deliberate, vote, and

determine an outcome). Availability means information is there to be learned if an individual wants to; accessibility implies that information has been learned and can be brought to mind; and applicability suggests that information can inform the individual's action. Several aspects of this definition are worth highlighting:

1. *Culture is information.* Some theorists (e.g., Herskovits, 1948; Triandis, 1994) have defined culture as the human-made part of the environment. We acknowledge that the human-made part of the environment (largely consisting of artefacts) *carries* or *embodies* cultural information, and cultural information enables humans to construct it; however, we distinguish what represents information from information itself.
2. *Culture is not a group.* Culture is often equated with a group or a collection of human individuals that continue to exist over a period of time where a group can vary in size from relatively small communities to the whole of humanity. We acknowledge that a culture is definable when a group is defined; however, we regard culture as information carried by members of a group or embodied in the artefacts constructed by them, and as such cultural information can continue to exist and inform human action even after the group ceases to exist (e.g., Classic Egyptian, Mesopotamian, Greek, Roman, Indian, Chinese cultures).
3. *Cultural information differs from genetic information in mode of transmission.* Culture consists of information that is socially transmitted, rather than genetically transmitted, from one person to another. Social transmission of cultural information can occur in various forms: formal schooling and explicit instruction, co-participation in joint activities, with or without language or other symbolic means. Although

genetic transmission occurs only from a biological parent to a biological child, cultural transmission occurs within and between generations, from older to younger, but also from younger to older, generations.

In this perspective, *cultural dynamics* is concerned with the trajectory of persistence and change of a set of cultural information and associated characteristics (e.g., structure, organization, and frequency and spatial distributions) in a group over time, and the social psychological mechanisms that drive these temporal dynamics.

**Sources of cultural dynamics.** Critical aspects of cultural dynamics are the *generation* and *retention* of information in the culture of a group. That is, when new cultural information is generated, it is added to the set; when it is retained over time (e.g., across generations), it is kept in the group's culture and potentially used in the future. The mechanisms that drive the generation and retention of cultural information is the *social transmission of cultural information* between people (see Kashima, in press-a, for recent research on this topic). This perspective is broadly in line with a host of theories of cultural evolution (e.g., Boyd & Richerson, 1985; Campbell, 1975; Cavalli-Sforza & Feldman, 1981; Dawkins, 1976; Richerson & Boyd, 2005; Sperber, 1996). Kashima, Peters, and Whelan (2008) called this metatheory *neo-diffusionism*.

In this perspective, there are at least four sources of cultural dynamics (Kashima, 2014). New information may be generated by *invention* within the group (e.g., Campbell, 1960; Simonton, 2010) or by *importation* from another group (e.g., Bartlett, 1923, 1932); once cultural information is present, it may be retained by *selection* – selecting it *in* for its benefit or selecting it *out* for its cost (e.g., Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981) – or by *drift*, non-selective stochastic transmission and retention (e.g., Bentley, Hahn,

& Shennan, 2004; Hahn & Bentley, 2003). The four sources of cultural dynamics differ in whether new information is added to a given culture, and whether adaptation is involved (i.e., benefit enhancement or cost reduction; see Table 1). Presumably, information is imported or selected during the process of adaptation; it may be invented or may drift due to stochastic processes that are neutral to adaptation (i.e., do not necessarily result in adaptation or maladaptation).

**Table 1. Sources of cultural dynamics**

	Random	Adaptive
Addition	Invention	Importation
Retention	Drift	Selection

**Note:** Each source differs in terms of whether it pertains to the addition of new cultural information or the retention (or removal) of cultural information within a population, and whether it occurs randomly or due to adaptive processes. We chose the term, random, here, but it may be better to say it is neutral (i.e., non-adaptive) or blind (i.e., without foresight). The characterization of invention as blind is due to Campbell (1960).

The sources of cultural dynamics presented here provide a framework for our discussions about diverse computational approaches to cultural dynamics. In this chapter, we begin our coverage with social psychological models of cultural dynamics (Abelson & Bernstein, 1963b; A. Nowak, Szamrej, & Latané, 1990), and then move to two prominent approaches to cultural dynamics – Axelrod’s (1997) model of cultural dissemination and evolutionary game theoretic approaches to evolution of cooperation. We show that these approaches focus on complementary aspects of cultural dynamics, and that each has unique strengths in dealing with some aspects, but not others. On one hand, the Axelrod model has been used to explore the dynamics deriving from transmissions of cultural information and the role of drift and to some extent of importation; however, it does not address invention, or most importantly, selection. On the other hand, evolutionary game theoretic approaches

have a unique strength in examining the importance of the selection process in cultural (and genetic) evolution. In this chapter, we will thus discuss how the existing approaches complement each other, and also point to the gap in the existing theory – neither has addressed the process of invention (other than as a random process of mutation).

We hasten to add that there are many analytical (for a systematic exposition, see McElreath & Boyd, 2007) and simulation (e.g., Carley, 1991; Hutchins, 1995) models of culture and cultural dynamics; however, they are too extensive and varied to be covered in the present chapter. It suffices to say that these attempts model some of the above processes at the micro-level of cognition and communication, and explore their consequences at macro-levels of organization, society, and culture. Nevertheless, there is a tradition in social psychology to attempt to describe the trajectory of cultural change with computer simulations of individuals' micro-level interactions. We start with this social psychological literature.

### **Social Psychology of Cultural Dynamics**

The social psychological tradition of cultural dynamics research began with public opinion dynamics. Note that in the current view of culture, a cultural element may be a public opinion. Given a certain proposition (i.e., cultural information) such as “Climate change is occurring, and it is largely human caused,” individuals in a group (e.g., Americans) can have a variety of opinions, and the frequency distribution of opinions on a binary variable such as pro vs. con, or a continuous bipolar dimension that indicates a degree of agreement or disagreement with the proposition can be regarded as an aspect of this group's culture. Obviously, the distribution of opinions about public issues such as climate change dynamically changes over time (e.g., Brulle, Carmichael, & Jenkins, 2012). Describing

the temporal trajectory of public opinions and examining the social psychological mechanisms underlying such opinion changes are both integral aspects of research on cultural dynamics (e.g., Kashima, 2014).

Abelson and Bernstein (1963a) were the first to examine such public opinion dynamics in computer simulations. In their model, agents had opinions about fluoridation of water supplies, interacted with other agents as a function of their relationships with them, conversed with them about fluoridation as a function of their interest in the issue, and updated their opinions and their relationships with their interaction partners; these steps occurred iteratively over time. Although most of the then existing models of attitude formation and change predicted a uniformly pro or con distribution of opinions in the community (Abelson, 1964), Abelson and Bernstein's simulation model showed a polarization of opinions – opinions would become more extreme – so that those agents that held initially pro opinions are likely to hold more pro opinions, and those that held con opinions become more con in opposition, as the controversy runs its course. In other words, public opinions could polarize and the polarization may persist and exacerbate, a finding that accorded much better with the empirical reality of controversies in realpolitik.

Although Abelson and Bernstein (1963a) showed that the emergence of consensus is not a universal consequence of simulation models, thus laying a foundation for simulating cultural dynamics, their model did not have an explicit spatial arrangement. It was A. Nowak et al. (1990) that explicitly defined a space (depending on its interpretation, it can be construed as a physical or social space). They suggested that agents can be located in a grid-like structure (i.e., lattice), where they would interact with other agents in their spatial neighbourhoods, exert social influences on each other as specified by Latané's (1981) social impact theory. More importantly, because their simulation model had an explicit definition

of space, they could also model the effect of spatial distance – it is known that opinions of others who are psychologically close (Latané's immediacy; also see Latané, Liu, Nowak, Bonevento, & Zheng, 1995, for empirical evidence) have a greater persuasiveness. Together with the nonlinearity in attitude change processes assumed in the model (see Lewenstein, Nowak, & Latané, 1992), their model produced spatial clustering of public opinions – even if there is a strong majority (i.e., a large proportion of agents has pro or con opinions), an opinion minority could survive if its members are spatially close to each other and cluster together in space (see Latané & L'Herrou, 1996, for empirical evidence).

These developments have enabled modeling of a single cultural element (e.g., opinion), and examination of the consequences of micro-level communication dynamics among agents on the macro-level trajectory of the opinion distribution in the group. This line of investigation can be extended in two different ways: multidimensionality and adaptiveness of culture. We will examine them in turn.

### **Multidimensionality of Culture: The Axelrod Model of Cultural Dissemination**

One of the characteristics of the foregoing model is that it models only a single cultural element. However, culture is obviously not a single element, but typically consists of multiple elements. Furthermore, cultural elements are not independent of each other, but are considered to be “patterned” (e.g., Kroeber & Kluckhohn, 1952; Triandis, 1996) or to form a configuration. How can a pattern of cultural elements emerge in computer simulations? Latané (1996) suggested one potential mechanism for the emergence of a cultural pattern. That is, when multiple cultural elements exist, but when each cultural element is influenced by the same mechanism of social influence such as dynamic social impact theory, even in the absence of any inherent relations between them, there will be a

spatial clustering of each cultural element. However, by virtue of the spatial clustering of these cultural elements, the elements that happen to occur within a given spatial cluster become correlated just by happenstance. A. Nowak, de Raad, and Borkowski (2012) showed this to be the case in their computer simulation (see also DellaPosta, Shi, & Macy, 2015; Harton & Bullock, 2007, for a review).

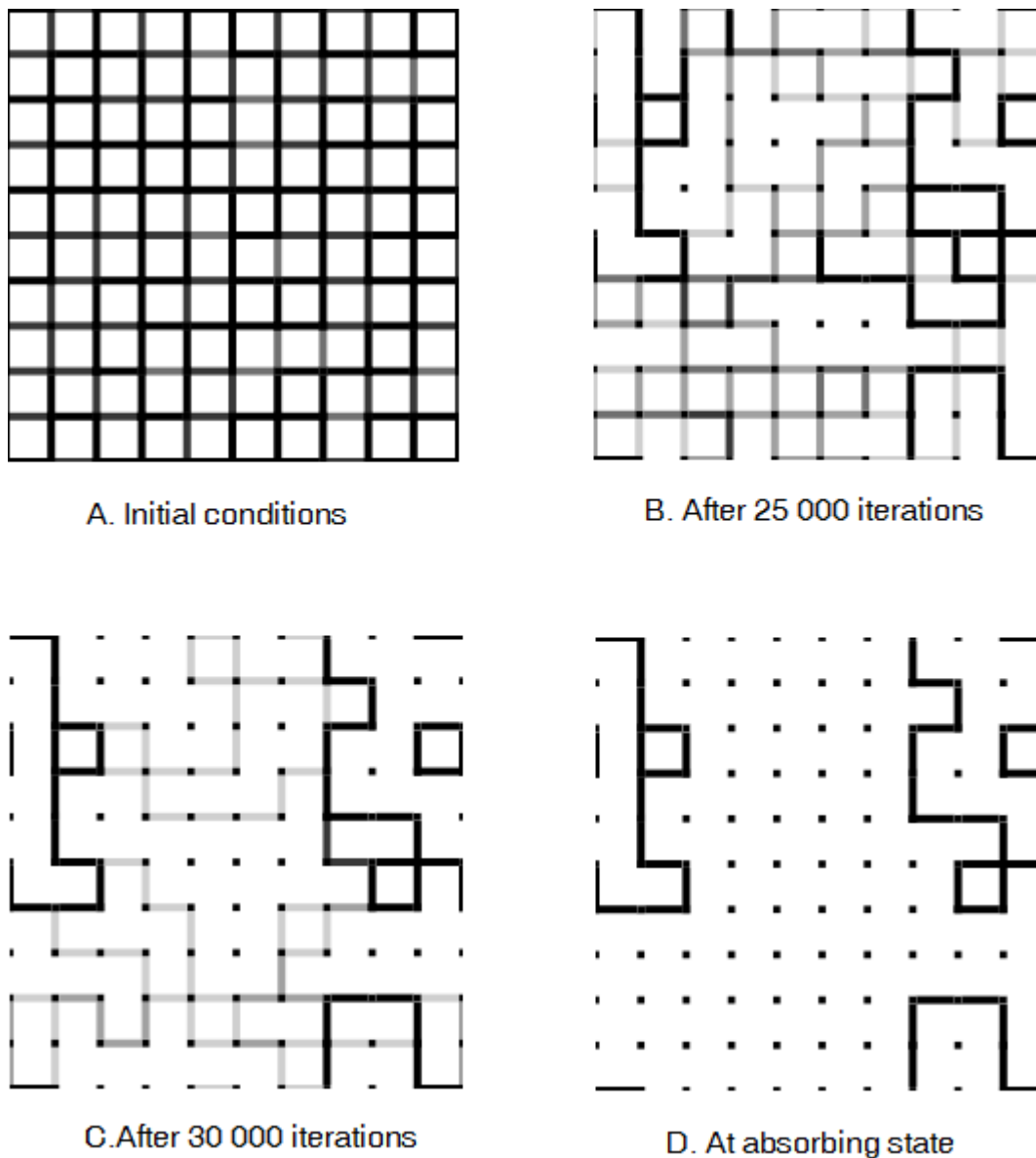
This line of investigation was further extended by the Axelrod model of cultural dissemination (Axelrod, 1997). In Axelrod's original model, cultural information is assumed to be represented as a set of multiple attributes (e.g., religion, language, taste) which each agent possesses. It is represented as a *culture vector* of  $F$  features, each of which can take  $q$  possible values (traits). For instance, one feature can be opinion about climate change, and then each agent can have one of several different types of opinions (e.g., "climate change doesn't exist," "climate change is happening, but not anthropogenic"). Another feature may be opinion about economic development (e.g., "market economy is the best form of resource distribution," "government intervention is sometimes important to regulate the economic process"). This way, there may be a number of cultural elements. When two agents interact, one agent's cultural information is transmitted to the other, and as a result changes the latter's culture vector. This transmission process, however, is constrained, so that the model assumes that cultural transmission occurs only when the agents already share some cultural information. In other words, if two agents have no cultural information in common, they cannot interact or transmit cultural information. This assumption can be interpreted as suggesting that (1) similar people tend to interact with each other (i.e., homophily), (2) people use their shared culture to communicate new cultural information, or both (1) and (2). Note that this process implies that cultural similarity begets cultural similarity, that is, culturally similar agents become more culturally similar to each other.

Does this mean that cultural differences eventually disappear? Even with multidimensional cultures, Axelrod (1997) showed that cultural diversity can persist under some circumstances.

### **Basic Model**

The dynamics of the model are as follows. Like A. Nowak et al. (1990), each agent is placed on a lattice site (every site on the lattice is occupied) with four neighbors (i.e., above, below, right, and left; von Neumann neighborhood), and only neighboring agents can interact. Initially, each agent is assigned a culture vector uniformly at random (that is, each of the  $F$  traits is given one of the  $q$  possible values at random). Thereafter, at each step, a random agent is chosen as the focal agent, and one of its neighbors is also chosen at random. The probability of the two agents interacting and influencing each other is determined by their cultural similarity, which is defined as the number of matching features between their culture vectors, that is, the number of corresponding features in which the two agents have the same trait value. When they interact, a random feature on which the two agents' values differ (if there is one) on the focal agent is changed to the value of the corresponding feature on the other agent.

This process is repeated until no more changes are possible, because any two neighboring agents either have identical culture vectors, or completely different culture vectors (no features in common, so they cannot interact). At this point, when no more change is possible, which is known as the *absorbing state*, there are two possible results. There is either a monocultural state, in which all agents have the same culture vector, or a multicultural state, in which any two agents in a contiguous region have the same culture, but neighboring agents on the region boundaries have completely different cultures (they can no longer interact as they have no features in common).

**Figure 1**

**Note:** Example of an Axelrod model run with number of features  $F = 5$ , number of traits  $q = 15$ , on a  $10 \times 10$  lattice. The darkness of the lines between sites indicates the cultural similarity (white is identical, black is completely different).

Figure 1 shows an example run of an Axelrod model using a visualization similar to that originally presented in Axelrod (1997), where the similarity of adjacent sites on the lattice is represented by the darkness of the line between the two sites. In the initial conditions (Figure 1A), adjacent sites are mostly completely different, although some have some small degree of similarity. In the intermediate stages (Figure 1B and Figure 1C),

cultural regions have started to form, and the boundaries between some cultural regions are colored light gray indicating they are quite similar. In the absorbing state (Figure 1D), which is a multicultural absorbing state which was reached after approximately 58 000 iterations, there are 5 cultural regions, the largest of which covers 75% of the lattice. Note that in this implementation of the model (Weaver, 2010), implemented in NetLogo (Wilensky, 1999) the lattice “wraps around” so that all sites on the lattice have exactly four neighbors, with no special cases on the edges of the lattice, unlike the original model specified by Axelrod (1997).

Whether a monocultural or multicultural absorbing state is reached, and if a multicultural absorbing state is reached, how many cultural regions exist, is influenced by a number of factors. Increasing  $q$ , the number of possible traits of each feature, results in an increased number of cultural regions, but in a nonlinear manner. All else being equal, when  $q$  is below a critical value, a monocultural absorbing state is the result. Increasing  $q$  above the critical value results in a multicultural absorbing state, with an increasing number of regions as  $q$  increases. Increasing the size of the neighborhood of agents with which the focal agent can interact decreases the number of cultural regions. Surprisingly, however, increasing the number of features,  $F$ , decreases the number of cultural regions, as does increasing the size of the lattice (which in the original Axelrod model also increases the number of agents, each lattice site being occupied) after a certain point.

So even this simple model demonstrates complex behavior, showing how homophily and social influence alone, in which two agents can only ever become more similar, does not necessarily result in a monoculture. In the language of statistical physics there is a nonequilibrium phase transition at the critical value of  $q$ , separating the ordered

(monocultural) phase from the disordered (multicultural) phase (Castellano, Marsili, & Vespignani, 2000).

### **Extensions of the Axelrod Model**

The original Axelrod model describes cultural transmission in dyadic social interactions in a grid-like social space with culture vectors taken from a fixed cultural space ( $F \times q$  matrix). However, it has been extended to investigate the implications of non-dyadic social interactions in more complex social and cultural structures.

**Non-dyadic Social Interactions.** The effect of *mass media* or other external cultural influences can be modeled as an external field acting on the culture vectors in the Axelrod model, causing a feature to become more similar to the external cultural message vector with a certain probability (Gandica, Charmell, Villegas-Febres, & Bonalde, 2011; González-Avella, Cosenza, Eguíluz, & San Miguel, 2010; González-Avella, Cosenza, & Tucci, 2005; González-Avella et al., 2006). Such an external cultural influence can, counter intuitively, lead to multicultural states when they would otherwise not have existed (González-Avella et al., 2005; González-Avella et al., 2006). A different, earlier, model of the mass media effect was a “generalized other”, a hypothetical extra neighbor agent with the most preferred value of each feature (Shibanai, Yasuno, & Ishiguro, 2001). In this model, rather than mass media being an external cultural influence, it is viewed as a social norm constructed from the culture vectors of all agents in the model.

The concept of *multilateral social influence* (that is, more than two agents interacting) was explicitly introduced by Parisi, Cecconi, and Natale (2003), although it was implicit in the “generalized other” of Shibanai et al. (2001). Parisi et al. (2003) show that, in the absence of homophily, stable multicultural states (neither monocultural nor anomic) can be sustained even in the presence of noise when social influence is multilateral. Kuperman

(2006) extends the Axelrod model by having an agent adopt a neighbor's trait only when it will then become similar to more of its neighbors. Extensions of the Axelrod model which incorporate multilateral social influence are thoroughly examined by Flache and Macy (2011), who show that such stable multicultural states can be sustained for wide ranges of noise levels even when homophily is assumed, as long as social influence is multilateral and not just dyadic. Rodríguez and Moreno (2010) incorporate both multilateral social influence and a mass media effect, again finding that multilateral social influence can maintain diversity in the presence of noise.

**Social and Cultural Spaces.** Rather than the standard lattice structure, cultural dynamics on more complex *social networks* have been investigated. One of them is a small-world network, which consists of densely connected clusters with sparse links between the clusters. Klemm, Eguíluz, Toral, and San Miguel (2003) find that the critical value of  $q$  increases with increasing density of long-distance connections. Guerra, Poncela, Gómez-Gardeñes, Latora, and Moreno (2010) explore the cultural dynamics on scale-free graphs (i.e., degree distribution follows a power law), which can be constructed by a process of preferential attachment: each new node is added to the graph by choosing an existing node to which to connect it with probability proportional to the existing node's degree (Barabási & Albert, 1999). Although these models have static social networks on which cultural dynamics unfold, more recent models have the social network co-evolve with culture (Centola, González-Avella, Eguíluz, & San Miguel, 2007; Vazquez, González-Avella, Eguíluz, & San Miguel, 2007), or allow for empty sites on the lattice and migration of agents (Gracia-Lázaro, Lafuerza, Floría, & Moreno, 2009), or both (Pfau, Kirley, & Kashima, 2013).

The cultural space in the original Axelrod model is simple and unstructured – a  $F \times q$  matrix. Most of the subsequent models used a uniform random distribution of culture

vectors as the initial conditions. As noted by Pace and Prado (2014), however, our understanding of the Axelrod model depends on this assumption, and different initial conditions can lead to very different behavior. An important exception to this general assumption is a model in which empirical opinion data is used as the initial culture vectors, which are found to have an approximately ultrametric distribution, leading to increased diversity at the absorbing state compared to random initial conditions (Valori, Picciolo, Allansdottir, & Garlaschelli, 2012). The effects of ultrametricity and other statistical properties of the initial culture vectors was further explored in Stivala, Robins, Kashima, and Kirley (2014). They found that when the agents have culture vectors generated by adding random noises to prototype culture vectors, they showed cultural dynamics similar to those generated by real opinion data, suggesting that the structure of cultural space needs to be further explored.

Nonetheless, these models still assume that different cultural features can change independently from each other. To the best of our knowledge no research has investigated the cultural dynamics when cultural representations are truly configural, so that a change in one cultural element may precipitate a change in other elements.

### **Sources of Cultural Dynamics in the Axelrod Model**

Some extensions of the Axelrod model addressed the cultural dynamics stemming from one of the sources of cultural dynamics, *cultural drift* (e.g., De Sanctis & Galla, 2009; Grauwin & Jensen, 2012; Konstantin Klemm, Víctor M Eguíluz, Raúl Toral, & Maxi San Miguel, 2003; Klemm, Eguíluz, Toral, & San Miguel, 2005; Parisi et al., 2003). In the Axelrod model, random perturbations can be added in the culture vectors as a form of drift. When a drift rate is small, the monocultural absorbing state becomes more likely; however, for large noise rates, a state of “*anomie*” is more likely to ensue, in which stable cultural regions fail

to form (Centola et al., 2007; Flache & Macy, 2011; Mäs, Flache, & Helbing, 2010). A high drift rate in this context implies that cultural transmission is noisy or inaccurate, and the findings here imply that inaccurate cultural transmissions are likely to result in a monocultural state. Hence, the model suggests that cultural information needs to be transmitted fairly accurately in order to maintain cultural diversity.

Arguably, the investigation of external cultural influences reviewed above taps a source of cultural dynamics akin to importation although further extensions including social identity and the like may be possible (e.g., Yamamoto, 2015). However, this research tradition has not investigated other sources of cultural dynamics such as *cultural invention*. That is to say, as noted by Pace and Prado (2014), Axelrod models (without noise) lack a method for the creation of novelty, since traits can only ever become extinct, and not be created. A possible major extension of the Axelrod model would be a mechanism to introduce new features or new traits, and compare the behavior of trait change compared to the creation of a whole new trait or feature. Furthermore, an obvious gap in this literature is an investigation of selection processes in which cultural information is selected in or out due to its adaptiveness or a lack thereof. A different tradition of research – evolution of cooperation – has addressed selection as its central focus, to which we now turn.

### **Adaptiveness: Evolution of Cooperation**

In the models of opinion dynamics and cultural dissemination discussed in the preceding section, agents interact with each other and transmit cultural information as a matter of course regardless of the social and psychological consequences of doing so. Nonetheless, as we noted earlier, one of the sources of cultural dynamics is *selection* (Table 1). That is, the more adaptive cultural information is (i.e., its retention, transmission, and

use are less costly and bring about more benefits at the psychological, social, and practical levels), the more likely it is to be selected and retained in the group's culture (e.g., Kashima, 2014; Kashima, in press-a, In press-b). However, the models of opinion dynamics and cultural dissemination do not take the consequences of cultural transmission into consideration in describing cultural dynamics.

In contrast, the literature on the evolution of cooperation is the most developed research tradition that addresses *selection* as a source of cultural dynamics. To be sure, there is a long tradition in social psychology, which investigates collective action in social dilemmas (e.g., Dawes, 1980; Kollock, 1998; Parks, Joireman, & Van Lange, 2013; Van Lange, Joireman, Parks, & Van Dijk, 2013). A social dilemma is defined as a situation in which individually beneficial actions produce collectively costly consequences. However, research on the evolution of cooperation addresses a question somewhat different from the typical experimental social psychological question, "How do humans behave in social dilemmas?" Rather, it can be construed as asking, "Under what circumstances, can cooperation as a cultural practice or a genetic characteristic remain in the population of agents?" In so doing, this research tradition attempts to answer under what circumstances can the cultural practice of cooperation be adaptive, and therefore selected in to remain in human culture.

This section describes some of the key concepts and main findings in the literature to provide an introduction to this voluminous literature. In particular, we review the standard game theoretic treatment of Prisoners' Dilemma as a prototypical game in evolution of cooperation, introduce an evolutionary game theoretic approach, and describe some of the key findings in this literature.

### **Cooperation in Social Dilemma**

The standard mathematical approach used to investigate the evolution of cooperation is based on game theory (Von Neumann & Morgenstern, 1953), whose fundamental assumptions are those of rationality and strategic interactions. That is, a rational player takes into account all the available information in the situation (i.e., the game) and tries to maximize their expected payoff by identifying an optimal sequence of choices of strategies (e.g., cooperation or defection). One of the best known examples of this approach is the classic Prisoner's Dilemma (PD) game (Axelrod, 1984; Rapoport & Chammah, 1965), a type of symmetric two-player simultaneous nonzero-sum game. Each of the two agents (typically called *players*), player A and B, makes a choice once from two potential *strategies*: cooperation (C) or defection (D). There are four possible outcomes for the game for each player depending on the combination of what each player decides to do:

Table 2: Players A and B's Choices and Outcomes for A

A's Choice	B's Choice	A's Outcome
C	C	Reward of mutual cooperation (R)
C	D	Sucker's payoff (S)
D	C	Temptation to defect (T)
D	D	Punishment for mutual defection (P)

Note.  $T > R > P > S$ .

Typically, the outcomes are displayed in the form of a payoff matrix like below.

		<b>B</b>	
		<b>C</b>	<b>D</b>
<b>A</b>	<b>C</b>	R	S
	<b>D</b>	T	P

The Prisoner's Dilemma payoff matrix is defined by the constraint  $T > R > P > S$ .

Looking at this situation from player A's perspective, it is easy to see that D always brings a

better outcome than C. If B chooses C, the outcome of A choosing D (T) is better than choosing C (R); if B chooses D, the outcome of A choosing D (P) is better than choosing C (S). In other words, D dominates C. Each player then choosing D is what is known as a *Nash Equilibrium* – the state in which unilaterally altering one’s behavior does not improve one’s payoff (Gintis, 2009; M. A. Nowak, 2006a). Assuming that the players behave so as to increase benefits and reduce costs, this is an equilibrium, in that it is a stable state in which neither player is willing to change his or her behavior unilaterally – the system of interlocking behaviors would remain stable. Nevertheless, given the payoff matrix, both players choosing D (outcome P) brings an outcome worse than they both choosing C (outcome R). Therein lies a dilemma – a strategic choice that brings a better outcome for each individual locks them into a mutually worse outcome.

This situation can be generalized to the case in which a game involves more than two players. One such game is the Public Goods game (e.g., Dawes, 1980; Hardin, 1968; also see Camerer, 2003; Gintis, 2009). Here, each player can either cooperate or defect, but the payoff matrix is generalized to the n person situation where the payoffs for cooperators and defectors are a function of the numbers of cooperators and defectors in the group. A player is always better off defecting than cooperating, but payoff is higher if everyone cooperates than if everyone defects. This is the essence of a *social dilemma*, an individually beneficial action resulting in a collectively unpalatable outcome.

**Iterated Games.** So far, we have considered the social situation in which agents make a choice only once (*one-shot game*). However, when agents play a game repeatedly (*iterated game*), the strategy to defect does not necessarily generate the highest payoff. For instance, imagine a Prisoner’s Dilemma game where players A and B repeatedly play a PD as shown in Table 2. If they decide to take turns, so that A cooperates and B defects in the first

round, A defects and B cooperates in the second round, A cooperates and B defects in the third round, etc., every two rounds both players cumulate the payoff  $S+T$ . If  $S+T > 2R$ , then the strategy of taking turns to cooperate and defect can maximize their outcomes (so, for a repeated PD, there is usually another constraint,  $S+T < 2R$ ). Thus, in an Iterated Prisoner's Dilemma game, the strategy of always defecting is not necessarily the best option.

More generally, if the number of rounds of the game is unknown (or at the very least the probability of playing additional rounds of the game is high), there are a variety of *conditional cooperation* strategies whose outcomes are better than the always-defect strategy in the long run. As we will see later, Trivers (1971) was perhaps the first to describe these type of strategies as driving the evolution of cooperation within a Prisoner's Dilemma; Axelrod (1984) extended this line of thinking. Put more broadly, the game theoretic formulation enables us to examine the adaptiveness of these game playing strategies, whether the long-term consequences of adopting cooperative strategies are beneficial or costly for the players. The foregoing discussion amounts to saying that a game playing strategy to cooperate is better for both players than a defective one under some circumstances, even if the social situation that dictates the outcome of their enactment constitutes a difficult social dilemma to resolve.

### **Evolutionary Game Theory**

Evolutionary game theory (e.g., Axelrod, 1984; Axelrod & Hamilton, 1981; Maynard Smith, 1982; M. A. Nowak, 2006a) extends the foregoing idea of iterative game plays by considering a population of players and their selections of strategies. Some terminological clarification is in order here. Although it is the standard practice in evolutionary game theory to call different ways of playing a game "*strategies*," this seems to us to be a reflection of the conceptual roots of this research tradition, the fact that game theory was

developed as a principled consideration of strategic moves when players are trying to outsmart each other. There is an unfortunate connotation that these game playing strategies are deliberate attempts to outperform each other. However, there is no reason why they must involve such deliberative foresights. In order to avoid the deliberative connotation, we will call strategies *genetic traits*, or *genetic tendencies* if they are genetically transmitted, whereas, we will call them *cultural traits*, *cultural practices*, or simply *practices* if such strategies are socially transmitted.

This terminological distinction should clarify what is being modeled. When a strategy like cooperation and defection is construed as a genetically transmitted trait, the resultant dynamics can be thought of as modeling biological evolution. However, when a strategy is interpreted as a cultural practice – a type of cultural information that describes a pattern of behavior – the resultant dynamics can be regarded as modeling cultural evolution. In either case, the payoff an individual receives corresponds to the notion of Darwinian *fitness* – average reproductive success in biological evolution and an average level of rewards in some sense in cultural evolution. The distribution of player's strategies defines a population state, which is formally equivalent to a notion of mixed strategy in game theory. Thus, evolutionary game theory investigates the emergent population dynamics and strategy distribution.

Now imagine that a large group of players play the Iterated Prisoner's Dilemma over multiple rounds of the game<sup>i</sup>. Furthermore, suppose that players can alter their strategy in response to the action played by the other players in previous rounds, such that a strategy that brings a better outcome is more likely retained. In any population of mixture of players that *always* choose to cooperate or defect (i.e., non-conditional co-operators or defectors), defection produces a higher average payoff than cooperation. Over time, cooperation tends

to disappear from the population and the relative proportion of defection increases as a result of the simulated selection process with all the group members eventually defecting – this happens even if there is only one defector, and the rest is all co-operators. In this sense, defection is an *evolutionarily stable strategy* (ESS)<sup>ii</sup>.

Nevertheless, as in the one-shot Prisoner's Dilemma Game, if all players choose to cooperate, everyone will be better off than if all players defect. By everyone cooperating, the entire group reaps benefit; by defecting, everyone will be worse off in the long run. But there is an incentive for each play to defect – the so-called *free rider problem*. If the Darwinian evolutionary process underlies the selection of genetically or socially transmittable behavioral patterns (e.g., cooperation and defection), how can cooperation evolve in a population? Thus, the evolution of cooperation is a theoretical puzzle in the Darwinian theory of biological and cultural evolution. And again, the answer is some form of conditional cooperation – a search for genetic or cultural mechanisms that enable conditional cooperative strategies to produce beneficial outcomes in the long run, even when the social situation that dictates the outcome of combinations of strategy choices involves social dilemmas like the Prisoner's or Public Goods Dilemma.

### **Mechanisms Promoting Cooperation**

The past theoretical research has identified a number of mechanisms that promote the evolution of cooperation. It is now well established that promotion and maintenance of cooperation within social dilemmas depends on the positive assortment of co-operators (S. A. West, Griffin, & Gardner, 2007). That is, there must be mechanisms or interaction structures to ensure that cooperators help other cooperators more than defectors. Below, we provide a brief overview of the five key mechanisms described by Nowak (M. A. Nowak, 2006b; Rand & Nowak, 2013) as well as others, which generally promote positive

assortment. M. A. Nowak (2006a) and McElreath and Boyd (2007) provide more formal treatment of the topics.

**Kin Selection.** Kin selection (or inclusive fitness) is a mechanism for the evolution of cooperation that arises if agents use conditional strategies based on kin relationship. This mechanism is based on the idea that cooperative behavior can emerge where the donor (the agent that benefits another agent) and the recipient (the agent that receives the benefit) are genetically related in the sense that there is a high probability of sharing a gene. That is, an individual is more likely to cooperate with closer relatives as compared to distant relatives or strangers. This tends to increase the average fitness of those that carry the genetic information that tends to produce cooperation among kinds. Hamilton's rule is typically used as a formalization for this mechanism (Hamilton, 1964a): the coefficient of relatedness,  $r$ , must exceed the cost-to-benefit ratio of the act for a cooperative action to be played. Although this may be based on genetic evolution, it can also be a result of cultural evolution.

**Direct Reciprocity.** Reciprocity is a mechanism by which agents use information about a history of agents' past behaviors to predict the probability of their cooperation. Direct reciprocity involves iterated encounters where agents play a game repeatedly across a finite number of rounds. Such repeated encounters between the same agents allow for reciprocation of cooperation, that is, when one cooperates, the other can return the favor by cooperating. As we noted earlier, Trivers (1971), Axelrod and Hamilton (1981), and Axelrod (1984) advanced direct reciprocity as a key mechanism for evolution of cooperation.

A well-known example of this type of strategy is 'tit-for-tat' developed by Anatol Rapoport, where a player cooperates in the first round and from then on always repeats whatever the other player did in the previous round. When it competed with other

strategies in a tournament of strategies, it was the best performing strategy in that it cumulated the highest payoffs of all strategies (Axelrod, 1984). Nonetheless, this strategy is not very tolerant of a mistake by another player – if players defect by mistake (by rare weakness of will, or a momentary lapse of judgment), strict tit-for-tat strategists descend to all defection. More flexible and tolerant strategies such as ‘generous tit-for-tat’ (one defection is generously tolerated) and ‘win–stay, lose–shift,’ have been shown to be effective once cooperators emerge in the population (M. A. Nowak, 2006a). Arguably, these are all variants of *conditional cooperation*, in which an agent cooperates if other agents cooperate. Indeed, approximately 50% of human players in experimental public goods games appear to be conditional co-operators (e.g., Fischbacher, Gächter, & Fehr, 2001).

**Indirect Reciprocity.** Direct reciprocity relies on the firsthand information about the probability of others’ cooperation obtained by the direct observation of their past behaviors. In contrast, indirect reciprocity rests on the information about the probability of others’ cooperation obtained not from the direct observation of their past behavior, but typically second-hand information based on third parties that directly observed these others’ past behaviors (e.g., M. A. Nowak & Sigmund, 1998; Panchanathan & Boyd, 2004). This is basically predicting other agents’ behavior based on their *reputation*, which may derive from gossip or other mechanisms of reputation management (e.g., Michelin ratings, university rankings). Agents can adopt conditional strategies and base their decision on reputation profiles of others in the population. For indirect reciprocity to be an effective mechanism in the evolution of cooperation, it requires that reputation information is as accurate as direct information obtaining from personal experience (e.g., M. A. Nowak & Sigmund, 1998; Panchanathan & Boyd, 2004).

It is important to note that, in order for indirect reciprocity to work, agents must have both cognitive and communicative capacities to remember their own interactions and monitor the ever-changing social network within their group, and to use language or other symbolic means to communicate reputational information, i.e., to gossip. Some have argued that this was one of the evolutionary bases of human language (Dunbar, 1998). In addition, stereotypes can serve as a basis of reputational information. If a group of agents is seen to be warm, communal, or moral (e.g., Eagly & Kite, 1987; Fiske, Cuddy, Xu, & Glick, 2002; Leach, Ellemers, & Barreto, 2007; Wojciszke, 2005a, 2005b), its members are likely inferred to be at least conditional co-operators. In particular, an agent's ingroup is often stereotyped as trustworthy (e.g., Brewer, 1979), and this autostereotype may be the basis of an ingroup favoritism in cooperation (e.g., Balliet, Wu, & De Dreu, 2014; Yamagishi & Kiyonari, 2000). Of course, this is not to say that stereotypes are accurate; however, it does present a new perspective on the evolution of stereotypes as a reputational mechanism.

**Network Reciprocity.** Human interactions are not random, but are typically structured in social networks – some individuals interact with each other more often than others. A social network is represented as a graph structure in which its vertices are occupied by agents, and its edges determine who interacts with whom. A payoff structure can be defined for each interaction, and repeated interactions between agents determine the overall outcome of each agent. Despite the complexity of the dynamics (see Perc, Gómez-Gardeñes, Szolnoki, Floría, & Moreno, 2013; Perc & Szolnoki, 2010; Szabó & Fáth, 2007, for an overview), the underlying principle in network reciprocity is the notion that ‘neighbors help each other’. Here, network reciprocity can promote the evolution of cooperation, because cooperators form clusters within which they cooperate with each other, which can prevail against exploitation by defectors. Recent work combining

evolutionary dynamics of group interactions on structured populations provides important insights into our understanding of the evolution of cooperation. A range of extensions can also be considered. For example, the population structure can be dynamic, so that agents can use 'active linking,' where individuals can choose to break unproductive links and establish new ones (Pacheco, Traulsen, & Nowak, 2006; Rand, Arbesman, & Christakis, 2011), or form social network ties with others based on their reputations (Fu, Hauert, Nowak, & Wang, 2008).

**Group Selection.** The idea of group selection (also called multi-level selection) has a controversial background in evolutionary biology and is frequently misunderstood (Stuart A. West, El Mouden, & Gardner, 2011). Wilson and Wilson (2008) describe how natural selection acts not only on individuals but also on groups. Consider the situation where a population is subdivided into groups. Individuals interact with other members of their ingroup in an evolutionary game that determines their fitness. In a simple scenario, defectors dominate cooperators within groups. However, if groups compete with each other, groups of cooperators outcompete groups of defectors. A little more theoretically, if selection processes operate at the level of groups (i.e., those groups with higher levels of average fitness are more likely to produce offspring than those with lower fitness levels), groups that contain more co-operators are more likely to do better than fewer co-operators. Under such conditions, multi-level selection acts as a powerful mechanism for the evolution of cooperation, especially if there are many small groups and if the migration rate between groups is not too large (Traulsen & Nowak, 2006). On the surface, there appears to be many similarities between spatial selection and multi-level selection. However, they are quite distinct mechanisms. In the former case selection (competition) occurs between individuals.

In the second case there is competition between individuals and competition between groups.

**Signaling.** As we noted earlier, the above-mentioned mechanisms typically promote positive assortment, i.e., the likelihood that cooperation is met by cooperation, so that cooperative actions co-occur in interacting agents. An additional mechanism for generating the co-occurrence of cooperation is based on signaling (e.g, Gintis, Smith, & Bowles, 2001; Skyrms, 2004). If individuals who are to cooperate in the future can signal their future action to cooperate to each other, they can coordinate their interaction (i.e., choose each other) and avoid the negative consequences of interacting with defectors.

There are a variety of signs that can be used to signal cooperation if agents with a genetic or cultural trait to cooperate have (a) a certain recognizable sign, (b) a capacity to recognize the sign, and (c) the tendency to cooperate when they recognize that sign. One such sign may be physical resemblance in appearance as a sign of kinship (e.g., Hamilton, 1964a, 1964b). Nonetheless, any arbitrary sign – what Dawkins (1976) called a ‘green beard’ – may work as long as it satisfies the above mentioned three properties (e.g., Riolo, Cohen, & Axelrod, 2001). The displaying of this sign may be costly (Gintis et al., 2001) or cheap (Robson, 1990; Skyrms, 2004); either way, positive assortment can occur under some circumstances. However, if agents develop a capacity to “fake” the sign (i.e., any individual can display the ‘green beard’, but not cooperate with other ‘green beards’), these agents will have higher fitness than others that cannot, and the sign will eventually be decoupled from the trait for cooperation.

**Additional Mechanisms.** In addition to these mechanisms, a number of additional mechanisms have also captured the attention of both social and evolutionary biology researchers. One is *voluntary participation*, in which, in addition to the default actions of

cooperate or defect, a player is also provided with a third action of not playing at all. In multi-player games such as the Public Goods game, significant levels of cooperation emerge, often in dynamic oscillations (Hauert, De Monte, Hofbauer, & Sigmund, 2002). Under specific circumstance, *punishment* can promote cooperative behavior in social dilemmas (Boyd & Richerson, 1992; Fehr & Gächter, 2002; Sigmund, Hauert, & Nowak, 2001; however, see Ohtsuki, Iwasa, & Nowak, 2009). Typically, punishment is embedded in models based on underlying mechanisms such as indirect reciprocity, spatial selection, and multi-level selection. Punishment may be implemented as a second stage action with additional costs to an individual, activated when other individuals in the game opt for defection. However, this approach can be undermined by the proliferation of second order 'free-riders', who cooperate but do not punish defectors.

### **Sources of Cultural Dynamics in Evolution of Cooperation**

In this section, we have attempted to describe how evolutionary game theory can be used to understand the evolution of cooperation largely from the perspective of adaptation, more specifically in terms of selection. Assuming that there are some agents who adopt conditional cooperation, when mechanisms are in place to ensure assortment, so that cooperators are more likely to interact with each other than with defectors, the enactment of the cultural practice (or genetic trait) of cooperation is likely to bring about better consequences than the cultural practice of defection on the average in the long run. Those practices that bring about greater benefit with less cost are assumed to be more likely to be transmitted and retained by agents. Therefore, provided that conditional cooperation brings about better outcomes than defection, conditional cooperation becomes more prevalent within the group. A culture of cooperation may thus emerge because under certain conditions where assortment is possible, cooperation is more adaptive.

Note that the current treatment is rather sketchy, and there is a more nuanced treatment about social transmission of cultural information in this framework. For example, cultural transmission may be biased (e.g., conformity – agents may be more likely to learn a practice that is more prevalent in a group; prestige – agents may be more likely to learn a practice that is used by a more prestigious agent) or unbiased, defective or cooperative tendencies may be both culturally and genetically inherited, and so on (e.g., Boyd & Richerson, 1985; Chudek & Henrich, 2011; Henrich & McElreath, 2003). Nonetheless, it is fair to say that the strength of the evolutionary game theoretic approaches lies in its theoretical treatment of adaptation, especially, the selection process in cultural dynamics.

### **Conclusions**

Social psychology has a tradition of modeling cultural dynamics. Starting with the public opinion dynamics that involve a binary choice between pro and con in a population of agents without any spatial structure (Abelson & Bernstein, 1963a) to dynamic social impact theory within a spatial structure (A. Nowak et al., 1990) and beyond, there has been a steady increase in the complexity with which social psychology has theorized cultural dynamics.

Nevertheless, further advances in modeling cultural dynamics have occurred outside the traditional boundary of social psychology although there are intriguing developments in social psychology as well (e.g., Denrell & Le Mens, 2007). One of the research traditions extended the univariate representation of culture (i.e., single opinion) to a multivariate representation (Axelrod, 1997), and the other has brought the evolutionary game theoretic framework to incorporate the adaptiveness of culture as a significant driver of cultural dynamics (e.g., McElreath & Boyd, 2007; M. A. Nowak, 2006a). These prominent research traditions have highlighted complementary aspects of cultural dynamics: the process and

consequences of cultural dissemination and the process of adaptation by selection. The complementarity of these developments poses an intriguing question for further exploration. How can these two research traditions be integrated, so that multidimensional and configural cultural information can be represented and the adaptiveness of the cultural information can be investigated within the same framework?

Theoretical questions aside, there are a number of pressing questions for models of cultural dynamics. First of all, there is an issue of empirical validation of models. Although the early challenges were more about how to model empirically well-established phenomena of opinion polarization (Abelson & Bernstein, 1963a) and persistence of diverse opinions (A. Nowak et al., 1990), the more recent models are mathematically sophisticated, but are yet to be tested or validated (however, some empirical tests are underway; for a recent review, see Rand & Nowak, 2013). Second, from apparently intractable intergroup conflicts to climate change, there are a number of challenges to the contemporary world that require a transformation of contemporary cultures (Kashima, In press-b; Wilson, Hayes, Biglan, & Embry, 2014). How can the models of cultural dynamics be used to benefit the public discourse and deliberation on planning and policy development for our common future? These are some of the difficult questions for the modeling of cultural dynamics, which future research will need to tackle.

### **Acknowledgment**

The preparation of this paper was facilitated by a grant from the Australian Research Council (DP130100845) to YK, GR, and MK.

## References

- Abelson, R. P. (1964). Mathematical models of the distribution of attitudes under controversy. In N. Freriksen & H. Gulliksen (Eds.), *Contributions to mathematical psychology* (pp. 142-160). New York, NY: Holt, Rinehart & Winston.
- Abelson, R. P., & Bernstein, A. (1963a). A computer simulation model of community referendum controversies. *Public Opinion Quarterly*, *27*(1), 93-122. doi: 10.1086/267152
- Abelson, R. P., & Bernstein, A. (1963b). A computer simulation model of community referendum controversies. *Public Opinion Quarterly*, *27*(1).
- Axelrod, R. (1984). *The evolution of cooperation*. New York, NY: Basic Books.
- Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution*, *41*(2), 203-226. doi: 10.1177/0022002797041002001
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, *211*(4489), 1390-1396. doi: 10.1126/science.7466396
- Balliet, D., Wu, J., & De Dreu, C. K. W. (2014). Ingroup favoritism in cooperation: A meta-analysis. *Psychological Bulletin*, *140*(6), 1556-1581. doi: 10.1037/a0037737
- Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, *286*(5439), 509-512. doi: 10.1126/science.286.5439.509
- Bartlett, F. C. (1923). *Psychology and primitive culture*. Cambridge, UK: Cambridge University Press.
- Bartlett, F. c. (1932). *Remembering: A study in experimental and social psychology*. Cambridge, UK: Cambridge University Press.
- Bentley, R. A., Hahn, M. W., & Shennan, S. J. (2004). Random drift and culture change. *Proceedings of the Royal Society of London, Series B*, *271*(1547), 1443-1450. doi: 10.1098/rspb.2004.2746
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Boyd, R., & Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, *13*(3), 171-195. doi: 10.1016/0162-3095(92)90032-Y
- Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin*, *86*(2), 307-324. doi: 10.1037/0033-2909.86.2.307
- Brulle, R. J., Carmichael, J., & Jenkins, J. C. (2012). Shifting public opinion on climate change: An empirical assessment of factors influencing concern over climate change in the u.s., 2002-2010. *Climatic Change*, *114*(2), 169-188. doi: 10.1007/s10584-012-0403-y

- Camerer, C. F. (2003a). *Behavioral game theory: Experiments on strategic interaction*. Princeton, NJ: Princeton University Press.
- Camerer, C. F. (2003b). Behavioral studies of strategic thinking in games. *Trends in Cognitive Sciences*, 7(5), 225-231. doi: 10.1016/s1364-6613(03)00094-9
- Campbell, D. T. (1960). Blind variation and selective retentions in creative thought as in other knowledge processes. *Psychological Review*, 67(6), 380-400. doi: 10.1037/h0040373
- Campbell, D. T. (1975). On the conflicts between biological and social evolution and between psychology and moral tradition. *American Psychologist*, 30(12), 1103-1126. doi: 10.1037/0003-066X.30.12.1103
- Carley, K. M. (1991). A theory of group stability. *American Sociological Review*, 56(3), 331-354.
- Castellano, C., Marsili, M., & Vespignani, A. (2000). Nonequilibrium phase transition in a model for social influence. *Physical Review Letters*, 85(16), 3536-3539.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution*. Princeton, NJ: Princeton University Press.
- Centola, D., González-Avella, J. C., Eguíluz, V. M., & San Miguel, M. (2007). Homophily, cultural drift, and the co-evolution of cultural groups. *Journal of Conflict Resolution*, 51(6), 905-929.
- Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in Cognitive Sciences*, 15(5), 218-226. doi: 10.1016/j.tics.2011.03.003
- Dawes, R. M. (1980). Social dilemmas. *Annual Review of Psychology*, 31(1), 169-193. doi: 10.1146/annurev.ps.31.020180.001125
- Dawkins, R. (1976). *The selfish gene*. Oxford, UK: Oxford University Press.
- De Sanctis, L., & Galla, T. (2009). Effects of noise and confidence thresholds in nominal and metric axelrod dynamics of social influence. *Physical Review E*, 79, 046108.
- DellaPosta, D., Shi, Y., & Macy, M. (2015). Why do liberals drink lattes? *American Journal of Sociology*, 120(5), 1473-1511.
- Denrell, J., & Le Mens, G. (2007). Interdependent sampling and social influence. *Psychological Review*, 114(2), 398-422. doi: 10.1037/0033-295X.114.2.398
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, 6(5), 178-190. doi: 10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8
- Eagly, A. H., & Kite, M. E. (1987). Are stereotypes of nationalities applied to both women and men? *Journal of Personality and Social Psychology*, 53(3), 451-462. doi: 10.1037/0022-3514.53.3.451

- Fehr, E., & Gächter, S. (2002). Altruistic punishment in humans. *Nature*, *415*(6868), 137-140. doi: 10.1038/415137a
- Fischbacher, U., Gächter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters*, *71*(3), 397-404. doi: 10.1016/S0165-1765(01)00394-9
- Fiske, S. T., Cuddy, A. J. C., Xu, J., & Glick, P. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878-902. doi: 10.1037//0022-3514.82.6.878
- Flache, A., & Macy, M. W. (2011). Local convergence and global diversity: From interpersonal to social influence. *Journal of Conflict Resolution*, *55*(6), 970-995. doi: 10.1177/0022002711414371
- Fu, F., Hauert, C., Nowak, M. A., & Wang, L. (2008). Reputation-based partner choice promotes cooperation in social networks. *Physical Review E*, *78*(2), 026117. doi: 10.1103/PhysRevE.78.026117
- Gandica, Y., Charmell, A., Villegas-Febres, J., & Bonalde, I. (2011). Cluster-size entropy in the axelrod model of social influence: Small-world networks and mass media. *Physical Review E*, *84*(4), 046109. doi: 10.1103/PhysRevE.84.046109
- Gintis, H. (2009). *The bounds of reason: Game theory and the unification of the behavioral sciences*. Princeton, NJ: Princeton University Press.
- Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *J Theor Biol*, *213*(1), 103-119. doi: 10.1006/jtbi.2001.2406
- González-Avella, J. C., Cosenza, M. G., Eguíluz, V. M., & San Miguel, M. (2010). Spontaneous ordering against an external field in non-equilibrium systems. *New Journal of Physics*, *12*(January), 013010. doi: 10.1088/1367-2630/12/1/013010
- González-Avella, J. C., Cosenza, M. G., & Tucci, K. (2005). Nonequilibrium transition induced by mass media in a model for social influence. *Physical Review E*, *72*(6), 065102(R). doi: 10.1103/PhysRevE.72.065102
- González-Avella, J. C., Eguíluz, V. M., Cosenza, M. G., Klemm, K., Herrera, J. L., & San Miguel, M. (2006). Local versus global interactions in nonequilibrium transitions: A model of social dynamics. *Physical Review E*, *73*(4), 046119. doi: 10.1103/PhysRevE.73.046119
- Gracia-Lázaro, C., Lafuerza, L. F., Floría, L. M., & Moreno, Y. (2009). Residential segregation and cultural dissemination: An axelrod-schelling model. *Physical Review E*, *80*(4), 046123. doi: 10.1103/PhysRevE.80.046123

- Grauwin, S., & Jensen, P. (2012). Opinion group formation and dynamics: Structures that last from nonlasting entities. *Physical Review E*, *85*(6), 066113.
- Guerra, B., Poncela, J., Gómez-Gardeñes, J., Latora, V., & Moreno, Y. (2010). Dynamical organization towards consensus in the axelrod model on complex networks. *Physical Review E*, *81*(5), 056105.
- Hahn, M. W., & Bentley, R. A. (2003). Drift as a mechanism for cultural change: An example from baby names. *Proceedings of the Royal Society of London, Series B*, *270*(Suppl. 1), S120-S123. doi: 10.1098/rsbl.2003.0045
- Hamilton, W. D. (1964a). The genetical evolution of social behavior. I. *J Theor Biol*, *7*(1), 1-16. doi: 10.1016/0022-5193(64)90038-4
- Hamilton, W. D. (1964b). The genetical evolution of social behavior. II. *J Theor Biol*, *7*(1), 17-52. doi: 10.1016/0022-5193(64)90039-6
- Hardin, G. (1968). The tragedy of the commons. *Science*, *162*(3859), 1243-1248. doi: 10.1126/science.162.3859.1243
- Harton, H. C., & Bullock, M. (2007). Dynamic social impact: A theory of the origins and evolution of culture. *Social and Personality Psychology Compass*, *1*(1), 521-540. doi: 10.1111/j.1751-9004.2007.00022.x
- Hauert, C., De Monte, S., Hofbauer, J., & Sigmund, K. (2002). Volunteering as red queen mechanism for cooperation in public goods games. *Science*, *296*(5570), 1129-1132.
- Henrich, J., & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology: Issues, News, and Reviews*, *12*(3), 123-135. doi: 10.1002/evan.10110
- Herskovits, M. J. (1948). *Man and his works: The science of cultural anthropology*. New York, NY: Knopf Doubleday.
- Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
- Kashima, Y. (2000). Conceptions of culture and person for psychology. *Journal of Cross-Cultural Psychology*, *31*(1), 14-32. doi: 10.1177/0022022100031001003
- Kashima, Y. (2014). How can you capture cultural dynamics? *Frontiers in Psychology*, *5*(995), 1-16. doi: 10.3389/fpsyg.2014.00995
- Kashima, Y. (in press-a). Cultural dynamics. *Current Opinion in Psychology*.
- Kashima, Y. (In press-b). Culture and psychology in the 21st century: Conceptions of culture and person revised. *Journal of Cross-Cultural Psychology*.
- Kashima, Y., & Gelfand, M. J. (2012). History of culture and psychology. In A. Kruglanski & W. Stroebe (Eds.), *Handbook of the history of social psychology* (pp. 499-520). New York, NY: Psychology Press.

- Kashima, Y., Peters, K., & Whelan, J. (2008). Culture, narrative, and human agency. In R. M. Sorrentino & S. Yamaguchi (Eds.), *Handbook of motivation and cognition across cultures* (pp. 393-421). San Diego, CA: Academic Press.
- Klemm, K., Eguíluz, V. M., Toral, R., & San Miguel, M. (2003). Global culture: A noise-induced transition in finite systems. *Physical Review E*, *67*(4), 045101(R).
- Klemm, K., Eguíluz, V. M., Toral, R., & San Miguel, M. (2003). Nonequilibrium transitions in complex networks: A model of social interaction. *Physical Review E*, *67*(2). doi: 10.1103/PhysRevE.67.026120
- Klemm, K., Eguíluz, V. M., Toral, R., & San Miguel, M. (2005). Globalization, polarization and cultural drift. *Journal of Economic Dynamics and Control*, *29*(1), 321-334. doi: 10.1016/j.jedc.2003.08.005
- Kollock, P. (1998). Social dilemmas: The anatomy of cooperation. *Annual Review of Sociology*, *24*(1), 183-214. doi: 10.1146/annurev.soc.24.1.183
- Kroeber, A. L., & Kluckhohn, C. (1952). *Culture: A critical review of concepts and definitions*. New York, NY: Vintage Books.
- Kuperman, M. N. (2006). Cultural propagation on social networks. *Physical Review E*, *73*(4), 046139. doi: 10.1103/PhysRevE.73.046139
- Latané, B. (1981). The psychology of social impact. *American Psychologist*, *36*(4), 343-356.
- Latané, B. (1996). Dynamic social impact: The creation of culture by communication. *Journal of Communication*, *46*(4), 13-25.
- Latané, B., & L'Herrou, T. (1996). Spatial clustering in the conformity game: Dynamic social impact in electronic groups. *Journal of Personality and Social Psychology*, *70*(6), 1218-1230.
- Latané, B., Liu, J. H., Nowak, A., Bonevento, M., & Zheng, L. (1995). Distance matters: Physical space and social impact. *Personality and Social Psychology Bulletin*, *21*(8), 795-805. doi: 10.1177/0146167295218002
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. Competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, *93*(2), 234-249. doi: 10.1037/0022-3514.93.2.234
- Lewenstein, M., Nowak, A., & Latané, B. (1992). Statistical mechanics of social impact. *Physical Review A*, *45*(2), 763-776. doi: 10.1103/PhysRevA.45.763
- Markus, H. R., & Kitayama, S. (1991). Culture and the self: Implications for cognition, emotion, and motivation. *Psychological Review*, *98*(2), 224-253. doi: 10.1037/0033-295x.98.2.224
- Mäs, M., Flache, A., & Helbing, D. (2010). Individualization as driving force of clustering phenomena in humans. *PLoS Comput Biol*, *6*(10), e1000959.

- Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge, UK: Cambridge University Press.
- McElreath, R., & Boyd, R. (2007). *Mathematical models of social evolution: A guide for the perplexed*. Chicago, IL: University of Chicago Press.
- Nowak, A., de Raad, W., & Borkowski, W. (2012). Culture change: The perspective of dynamical minimalism. In M. J. Gelfand, C.-Y. Chiu & Y.-Y. Hong (Eds.), *Advances in culture and psychology* (Vol. 2, pp. 249-314). New York: Oxford University Press.
- Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitudes to public opinions: A dynamic theory of social impact. *Psychological Review*, *97*(3), 362-376.
- Nowak, A., & Vallacher, R. R. (1998). *Dynamical social psychology*. New York: Guilford.
- Nowak, M. A. (2006a). *Evolutionary dynamics: Exploring the equations of life*. Cambridge, MA: Harvard University Press.
- Nowak, M. A. (2006b). Five rules for the evolution of cooperation. *Science*, *314*(5805), 1560-1563. doi: 10.1126/science.1133755
- Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image scoring. *Nature*, *393*(6685), 573-577. doi: 10.1038/31225
- Ohtsuki, H., Iwasa, Y., & Nowak, M. A. (2009). Indirect reciprocity provides only a narrow margin of efficiency for costly punishment. *Nature*, *457*(7225), 79-82. doi: 10.1038/nature07601
- Pace, B., & Prado, C. P. (2014). Axelrod's model with surface tension. *Physical Review E*, *89*(6), 062804. doi: 10.1103/PhysRevE.89.062804
- Pacheco, J. M., Traulsen, A., & Nowak, M. A. (2006). Active linking in evolutionary games. *J Theor Biol*, *243*(3), 437-443. doi: 10.1016/j.jtbi.2006.06.027
- Panchanathan, K., & Boyd, R. (2004). Indirect reciprocity can stabilize cooperation without the second-order free rider problem. *Nature*, *432*(7016), 499-502. doi: 10.1038/nature02978
- Parisi, D., Cecconi, F., & Natale, F. (2003). Cultural change in spatial environments: The role of cultural assimilation and internal changes in cultures. *Journal of Conflict Resolution*, *47*(2), 163-179.
- Parks, C. D., Joireman, J., & Van Lange, P. A. M. (2013). Cooperation, trust, and antagonism: How public goods are promoted. *Psychological Science in the Public Interest*, *14*(3), 119-165. doi: 10.1177/1529100612474436
- Perc, M., Gómez-Gardeñes, J., Szolnoki, A., Floría, L. M., & Moreno, Y. (2013). Evolutionary dynamics of group interactions on structured populations: A review. *Journal of The Royal Society Interface*, *10*(80), 20120997. doi: 10.1098/rsif.2012.0997

- Perc, M., & Szolnoki, A. (2010). Coevolutionary games: A mini review. *Biosystems*, *99*(2), 109-125. doi: 10.1016/j.biosystems.2009.10.003
- Pfau, J., Kirley, M., & Kashima, Y. (2013). The co-evolution of cultures, social network communities, and agent locations in an extension of axelrod's model of cultural dissemination. *Physica A*, *392*(2), 381-391. doi: 10.1016/j.physa.2012.09.004
- Railsback, S. F., & Grimm, V. (2012). *Agent-based and individual-based modelling: A practical introduction*. Princeton, NJ: Princeton University Press.
- Rand, D. G., Arbesman, S., & Christakis, N. A. (2011). Dynamic social networks promote cooperation in experiments with humans. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(48), 19193-19198. doi: 10.1073/pnas.1108243108
- Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*(8), 413-425. doi: 10.1016/j.tics.2013.06.003
- Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor, MI: University of Michigan Press.
- Richerson, P. J., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution*. Chicago, IL: University of Chicago Press.
- Riolo, R. L., Cohen, M., & Axelrod, R. (2001). Evolution of cooperation without reciprocity. *Nature*, *414*(6862), 441-443.
- Robson, A. J. (1990). Efficiency in evolutionary games: Darwin, nash and the secret handshake. *J Theor Biol*, *144*(3), 379-396. doi: 10.1016/S0022-5193(05)80082-7
- Rodríguez, A. H., & Moreno, Y. (2010). Effects of mass media action on the axelrod model with social influence. *Physical Review E*, *82*(1), 016111. doi: 10.1103/PhysRevE.82.016111
- Shibanai, Y., Yasuno, S., & Ishiguro, I. (2001). Effects of global information feedback on diversity: Extensions to axelrod's adaptive culture model. *Journal of Conflict Resolution*, *45*(1), 80-96.
- Sigmund, K., Hauert, C., & Nowak, M. A. (2001). Reward and punishment. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(19), 10757-10762. doi: 10.1073/pnas.161155698
- Simonton, D. K. (2010). Creative thought as blind-variation and selective-retention: Combinatorial models of exceptional creativity. *Physics of Life Reviews*, *7*(2), 156-179. doi: 10.1016/j.plrev.2010.02.002
- Skyrms, B. (2004). *The stag hunt and the evolution of social structure*. Cambridge, United Kingdom: Cambridge University Press.
- Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Oxford, UK: Blackwell.

- Stivala, A., Robins, G., Kashima, Y., & Kirley, M. (2014). Ultrametric distribution of culture vectors in an extended Axelrod model of cultural dissemination. *Scientific Reports*, *4*, 4870. doi: 10.1038/srep04870
- Szabó, G., & Fáth, G. (2007). Evolutionary games on graphs. *Physics Reports*, *446*(4-6), 97-216. doi: 10.1016/j.physrep.2007.04.004
- Traulsen, A., & Nowak, M. A. (2006). Evolution of cooperation by multilevel selection. *Proceedings of the National Academy of Sciences*, *103*(29), 10952-10955. doi: 10.1073/pnas.0602530103
- Triandis, H. C. (1964). Cultural influences upon cognitive processes. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 1). New York: Academic Press.
- Triandis, H. C. (1989). The self and social behavior in differing cultural contexts. *Psychological Review*, *96*(3), 506-520. doi: 10.1037/0033-295X.96.3.506
- Triandis, H. C. (1994). *Culture and social behavior*. New York, NY: McGraw-Hill.
- Triandis, H. C. (1996). The psychological measurement of cultural syndromes. *American Psychologist*, *51*(4), 407-415. doi: 10.1037/0003-066X.51.4.407
- Trivers, R. L. (1971). The evolution of reciprocal altruism. *The Quarterly Review of Biology*, *46*(1), 35-57. doi: 10.1086/406755
- Valori, L., Picciolo, F., Allansdottir, A., & Garlaschelli, D. (2012). Reconciling long-term cultural diversity and short-term collective social behavior. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(4), 1068-1073. doi: 10.1073/pnas.1109514109
- Van Lange, P. A. M., Joireman, J., Parks, C. D., & Van Dijk, E. (2013). The psychology of social dilemmas: A review. *Organizational Behavior and Human Decision Processes*, *120*(2), 125-141. doi: 10.1016/j.obhdp.2012.11.003
- Vazquez, F., González-Avella, J. C., Eguíluz, V. M., & San Miguel, M. (2007). Time-scale competition leading to fragmentation and recombination transitions in the coevolution of network and states. *Physical Review E*, *76*(4), 046120.
- Von Neumann, J., & Morgenstern, O. (1953). *Theory of games and economic behavior* (3rd ed.). Princeton, NJ: Princeton University Press.
- Weaver, I. (2010). Dissemination of culture. *NetLogo user community models*.  
<http://ccl.northwestern.edu/netlogo/models/community/Dissemination%20of%20Culture>
- West, S. A., El Mouden, C., & Gardner, A. (2011). Sixteen common misconceptions about the evolution of cooperation in humans. *Evolution and Human Behavior*, *32*(4), 231-262. doi: 10.1016/j.evolhumbehav.2010.08.001
- West, S. A., Griffin, A. S., & Gardner, A. (2007). Evolutionary explanations for cooperation. *Curr Biol*, *17*(16), R661-672. doi: 10.1016/j.cub.2007.06.004

- Wilensky, U. (Producer). (1999). Netlogo. Retrieved from <http://ccl.northwestern.edu/netlogo/>
- Wilson, D. S., Hayes, S. C., Biglan, A., & Embry, D. D. (2014). Evolving the future: Toward a science of intentional change. *Behavioral and Brain Sciences*, 37(4), 395-416. doi: 10.1017/S0140525X13001593
- Wilson, D. S., & Wilson, E. O. (2008). Evolution "for the good of the group". *American Scientist*, 96(5), 380-389. doi: 10.1511/2008.74.1
- Wojciszke, B. (2005a). Affective concomitants of information on morality and competence. *European Psychologist*, 10(1), 60-70. doi: 10.1027/1016-9040.10.1.60
- Wojciszke, B. (2005b). Morality and competence in person- and self-perception. *European Review of Social Psychology*, 16(1), 155-188. doi: 10.1080/10463280500229619
- Yamagishi, T., & Kiyonari, T. (2000). The group as the container of generalized reciprocity. *Social Psychology Quarterly*, 63(2), 116-132. doi: 10.2307/2695887
- Yamamoto, K. (2015). Mobilization, flexibility of identity, and ethnic cleavage. *Journal of Artificial Societies and Social Simulation*, 18(2), 8. doi: 10.18564/jasss.2669

## Notes

---

<sup>i</sup> Typically, in the Iterated Prisoner's Dilemma game, an additional constraint on the payoff matrix values is employed,  $2R > (T + S)$ , so that players are not collectively better off if they simply alternate between playing cooperate and defect actions.

<sup>ii</sup> Here, the ESS constitutes a Nash equilibrium, but they are not strictly equivalent. For detailed discussions about the relationship between ESS and Nash equilibria, see M. A. Nowak (2006a).