Emergent cultural signatures and persistent diversity: A model of conformity and consistency

Jenna Bednar
University of Michigan, USA

Aaron Bramson
University of Michigan, USA

Andrea Jones-Rooy
University of Michigan, USA

Scott Page
University of Michigan, USA

Abstract

Empirical evidence demonstrates that cultures exist, they differ from one another, they're coherent and yet diversity persists within them. In this paper, we describe a multi-dimensional model of cultural formation that produces all of these properties. Our model includes two forces: an internal desire to be *consistent* and social pressure to *conform*. When both forces operate, the society converges to a coordinated behavior that is consistent across the attributes. We find that convergence in the two-force model is slower than a pure conformity model and that a preponderance of one force over the other slows convergence, rather than hastening it. We further find that the two forces amplify small errors in individual behavior and prove capable of producing substantial persistent diversity.

Corresponding author:
Jenna Bednar, Department of Political Science and Center for Political Studies, University of Michigan, 4259 ISR 426 Thompson Street, Ann Arbor, MI 48106, USA
Email: jbednar@umich.edu

Empirical research on cultural differences reveals four broadly accepted findings. First, cultures exhibit coherence in both actions and belief systems. These cultural signatures are socially, economically, and politically meaningful. Second, these actions and behaviors also exhibit consistency. This consistency allows policy makers to anticipate and predict responses based on cultural affiliations and it enables scholars to identify artifacts as belonging to particular groups. Third, despite the identifiability of consistent group-level signatures, cultures exhibit substantial within-group heterogeneity. Not all Germans are alike. Finally, and perhaps most importantly, cultures differ. These differences influence the management of economies, political systems, and the environment.

The formal literature on cultural differentiation and integration considers the first and last regularities, as Macy et al. describe them: 'the disproportionate homogeneity in social groups and the persistence of diversity across groups' (Macy et al., 2003: 2). Within-group homogeneity can be produced by mild incentives for coordination (Durrett and Levin, 2005; Kuran and Sandholm, 2008). These same models implicitly demonstrate inter-group diversity: geographically disjoint groups would be very unlikely to converge on identical ensembles of beliefs and behaviors. A more subtle logic based on preferential interaction can also explain cultural differences between groups. Axelrod's (1997) culture model assumes that agents mediate interactions based on social distances. As a result, people cluster into like-minded groups as thick boundaries emerge between internally homogeneous cultures (see also Centola et al., 2007; Friedkin, 1984; Friedkin and Johnsen, 1997; Kitts et al., 1999; Klemm et al., 2003; Marsden and Friedkin, 1993).[1]

The existing models provide a foundation from which to explore the causes of the other two findings: consistency across domains and persistent diversity within interacting populations. In this paper, we describe a model that has as its base a set of standard coordination games. The coordination game, in some form, underpins most culture models. We then add a behavioral assumption that individuals prefer to behave consistently across domains. We might expect behavior and beliefs to transfer across domains for either of two reasons. First, coherence may require less cognitive effort (Bednar and Page, 2007). Consistent behavior entails less cognitive dissonance (Festinger, 1957). Second, an individual's coherence may be informationally more efficient for others. If members of a culture assign types, such as high or low trust, to others, then they can update beliefs much more

easily and accurately if they assume those types apply across domains. We then show that this model can produce all four findings.

As a first step, we define what we mean by a consistent cultural signature.[2] Within a culture, actions on each domain can be coordinated, i.e. everyone can shake hands or bow, but this need not imply that behavior in one domain bears any resemblance to behaviors in other domains. Ethnographic research suggests that human cultures possess signatures, or what might be described as coherence across domains of actions and beliefs. Signatures allow actors within cultures to anticipate responses across wide settings of interaction. For example, if a culture values competitiveness in its economy, its schools may teach and reward competitiveness as well. If a culture is based upon a strict social hierarchy, employees may be more likely to encounter glass ceilings obstructing their merited promotion if they come from a class perceived as subordinate. In contrast, a culture may be technically speaking coherent, but lack identifiable linkages across domains, such as a society where children are the center of attention in social occasions, but that sets school funding as its lowest priority.

One can model individual attributes as a vector of traits, and a culture as the preponderance of similar vectors within a community. For example, a culture may coordinate on the vector (7, 4, 3, 5). If most of the people within a group exhibit these traits, then we can think of this as a *coherent culture*. For a culture to exhibit a *consistent signature*, we need the traits in one domain to be interpretable relative to traits in other domains. In other words, in the vector (7, 4, 3, 5), we would need the seven and the four to have an explicit relationship to one another. In the simplest version of our model, we assume that traits, what we call attributes, take binary values. We further assume that if two attributes take the same value, then they are consistent. Therefore, a culture that coordinated on the vector (1, 1, 1, 1) would be consistent, but one that coordinated on (0, 1, 0, 1) would not be.

This construction allows us to anticipate how the two cultures might respond to a new domain of interaction. We would expect that the members of the first culture would choose action 'one', but we could only guess at the behavior of the members of the second culture. Predictions of cultural consistency can and have been made empirically. Fisman and Miguel (2007) find a correlation between home country corruption levels and unpaid parking tickets issued to diplomats in New York City, and Miguel et al. (2008) have found a correlation between a soccer player's tendency be penalized and the prevalence of violence in his home country. In both cases, behavior translates across distinct domains demonstrating consistency.

Our construction exogenously imposes a single perceived meaning to each action. This does not allow for socially constructed meanings which are well

established by interpretive models of culture (Wedeen, 2002). Our current model can be extended to allow actions to have sets of possible meanings or to have multi-dimensional meanings. Either extension would allow for the social construction of the meaning of acts. They would also allow individuals to reinterpret or rationalize their actions to establish consistency. The effect of including socially constructed meanings would be to increase intra-cultural heterogeneity. Our approach, which fixes the meaning of acts, makes the production of heterogeneity more difficult, so it is the natural place to start.

By heterogeneity, we mean something different from the polarization of sub-cultures.[3] We mean the lack of convergence within an interacting group. The fact that existing models do not produce significant within-culture variation can hardly be called a failure given the purposes of those models. By adding a small noise or error term, coordination models produce proportional levels of heterogeneity, as we show. In our model, small amounts of noise or error become amplified producing substantial heterogeneity. The more dimensions to behavior and the more domains of interaction, the greater this amplification. Thus, our model suggests behavioral trembles as a candidate explanation for the observed substantial persistent heterogeneity of cultures.

Recall that our model differs from extant models in that it includes two incentives: a desire to coordinate and a desire to be consistent. To show how these two forces operate, we describe three models: a *pure conformity* model, a *pure consistency* model, and a *combined conformity/consistency* model. As we have no reason to believe that conformity and consistency matter equally, we also vary the weight across the two forces in the combined model. For each model we first omit the possibility of error, and then we include random errors and solve for equilibrium distributions of behavior. The error-free, single force models produce intuitive results: in the pure conformity model, individuals adopt identical but internally inconsistent attributes. Whatever cultural differences exist arise from randomness (the odds of coordinating on the same attributes are low) or from different initial conditions (the attributes that are most prevalent initially tend to become the dominant cultural attributes). This simple social conformity (i.e., coordination) model can therefore explain inter-cultural heterogeneity as distinct equilibria of the coordination dynamic, but it cannot produce intra-cultural heterogeneity. Moreover, these equilibria cannot be interpreted as possessing a cultural signature.[4] Similarly, the error-free, pure consistency model produces consistent individuals but no intra-culture consistency. That is, it produces internally consistent individuals, but social coordination only arises by chance. When we introduce both forces of consistency and conformity to the model, the error-free version produces consistent individuals and coherent cultures but no within-cultural heterogeneity. We find though

that the model takes a comparatively long time to converge, which suggests complex dynamics. Our formal time to convergence results provide an alternative method for interpreting the amplification of trembles, namely that the complex dynamics allow errors to accumulate.

1 Meaningful cultural signatures and the persistence of internal group diversity

Survey research across social science disciplines has consistently revealed substantial differences between cultures. In fact, this inter-cultural variation provides a foundation for nearly all social scientific comparative studies. The nature of area studies research implicitly assumes recognizable and significant differences between behaviors of peoples in different geographical regions, be they informal societies, communities, cities, or countries. Inglehart, in summarizing data from the World Values Survey, concludes that 'cultural variation is ... relatively constant within a given society, but shows relatively great variation between different societies' (Inglehart, 1997: 166).[5] The French, for example, tend to be more uncertainty averse than Americans (Hofstede, 2001), and Danish attitudes about well-being can be consistently distinguished from French, Italian, or Portuguese attitudes.

Complementing these survey results, psychological and economic experiments reveal systematic differences in patterns of thought across countries. Nisbett (2003) finds that citizens of Eastern countries take context into account far more than Westerners and as a result are much more willing to accept logical contradiction. Henrich et al., (2001) study fifteen small-scale societies across five cultures finding substantive evidence of inter-cultural behavioral variation. In another study Henrich (2000) finds that the economic behavior of Peruvian communities varies widely from the behavior of a Los Angeles control group, which suggests that 'economic reasoning may be heavily influenced by cultural differences – that is, by socially transmitted rules about how to behave in certain circumstances (economic or otherwise) that may vary from group to group as a consequence of different cultural evolutionary trajectories' (Henrich, 2000: 973).

Despite observable cultural signatures, people within cultures differ widely. Not all German people think and act the same way, nor do all members of the Itza' or the !Kung. In fact, the differences within cultures are as substantial as the differences between them (Au, 1999; Hofstede, 2001; Inglehart, 1997; Pelto and Pelto, 1975; Thompson, 1975). Au (1999) captures within-culture diversity related to work beliefs, finding that some countries that share similar cultural average scores can have wider or tighter spreads in the distribution of the population, and this variation may affect

cross-cultural comparisons as much as the average. A simple plot of two factor scores of two countries from the World Values Survey data provides striking demonstration of all three regularities: inter-cultural diversity, identifiable cultural signatures, and intra-cultural diversity.

We generated Figure 1 using data from the World Values Survey (Inglehart and Welzel, 2005). On the *x*-axis the Traditional–Secular-rational dimension captures how important religion is to respondents in each country. High scores on this dimension correspond with high values placed on concepts like family, tradition, and deference to authority. On the *y*-axis the Survival–Self-expression dimension captures the differences in responses on questions related to materialist or post-materialist values. High scores correspond with high values placed on concepts like diversity, imagination, tolerance, environmental protection, and involvement in political and economic life, as well as with high interpersonal trust. Not only do responses differ considerably between countries, but also responses from Sweden demonstrate greater within-culture diversity. These factors aggregate multiple responses from each individual, so were the variation due to random error, within-country variation would be very small.

Evidence of intra-cultural variation and our proposed explanation say nothing of its importance. Cultural, economic, and social behavior may differ in meaningful ways. Durham (1991) demonstrates variety in types of marriage custom within Tibetan culture. Thompson (1975) provides evidence of significant intra-cultural variation in willingness to accept delayed economic gratification between three communities in Uganda. A study of a series of six cultures across four continents by Minturn and Lambert (1964) and Whiting (1963) found that all but one variable on child rearing behavior was better captured by intra-cultural variation than by inter-cultural variation. In sum, within-cultural differences matter as much as cross-cultural differences, and for many of the same reasons.

2 Microfoundations: Conformity and consistency

The development of a meaningful cultural signature implies that individuals within a community conform their behavior to match one another's, and also that there is some relationship that ties their behaviors and beliefs together from one activity or domain to the next, creating consistency across behaviors. In this section we describe and then model the two forces of social conformity and internal consistency.

Social conformity, the inclination to become more like those around oneself, can be unpacked into four distinct individual-level desires and incentives: (1) the desire to fit in with others, (2) the strategic benefit from coordination, (3) the incentive to free ride on the information of others, and (4) the tendency
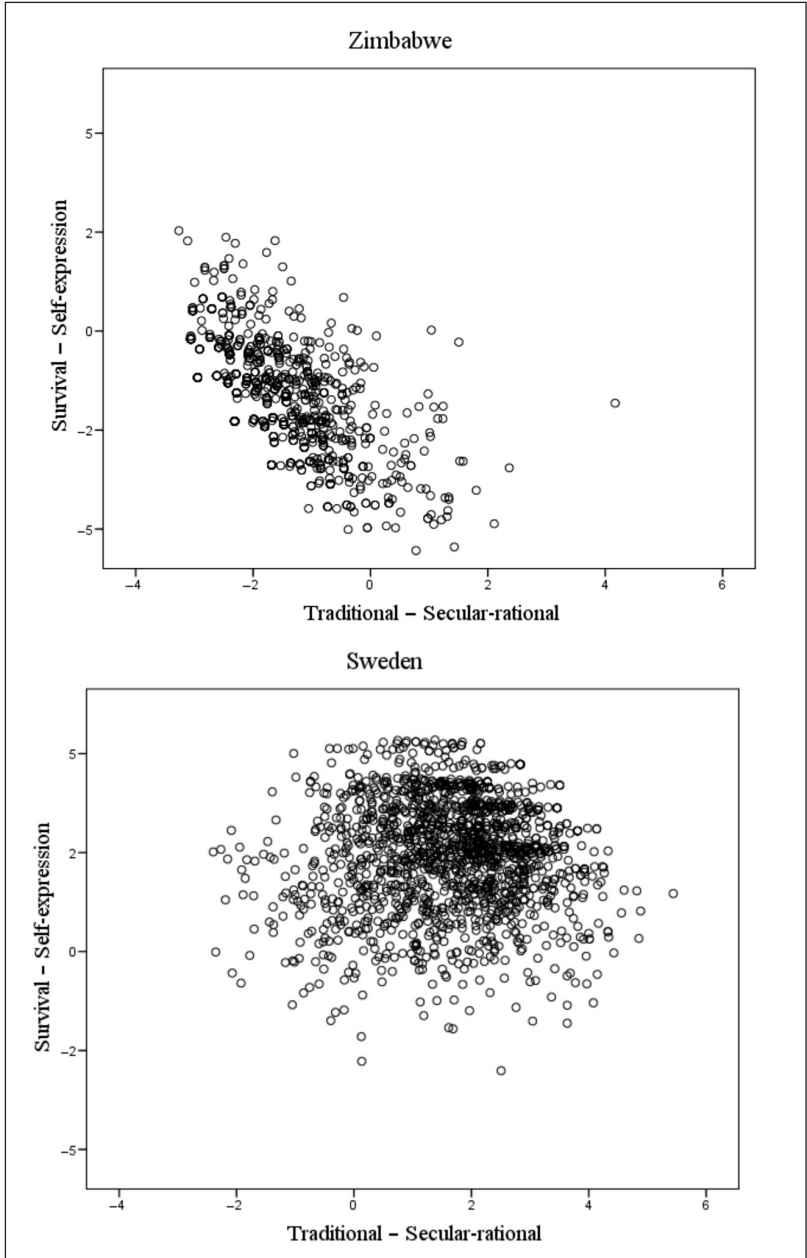
Figure 1. Respondent diversity within and between Zimbabwe and Sweden

to interact with people similar to oneself. It is a well established observation in social psychology that people tend to mimic the behaviors, beliefs, and attributes of those with whom they interact. Social pressure can also impart desire to fit in with others (Bernheim, 1994; Kuran, 1995). And, if others positively reinforce conforming behavior, then conformity itself can become a conditioned response (Pavlov, 1903; Skinner, 1974). Finally, people who interact frequently act similarly, dress similarly, reveal similar preferences (Axelrod 1997), and react similarly to novel situations (Simon, 1982).[6]

Although conformity has several well accepted causes, attempts to identify the extent of conformity run up against an identification problem (Brock and Durlauf, 2006). In addition to conforming, people also choose to be around those who act as they do – this is the observation of *homophily*.[7] Homophily curbs group mergers because people avoid interacting with others who are not like themselves. Adding social influence to models of homophily exacerbates these effects: when individuals interact with others like themselves, and also actively become more similar to them, polarization between groups is even more pronounced. In our model we take the interacting groups as fixed and rule out the possibility of subcultures of this sort, acknowledging that the possibility of subcultures would create further intra-cultural heterogeneity.

Conformity need not be divorced from incentives. Individuals often mimic selectively, looking only to the behaviors of their more successful neighbors (Kennedy, 1988). Moreover, common behavior need not imply conformity. People who face similar problems may construct similar solutions without imitating just as students who enroll in the same class and take identical exams may produce similar answers without copying. Seminal works in psychology by Pavlov (1903) and Skinner (1974) connect positive reinforcement and the conditioning of learned responses. Similar learning environments could, therefore, condition near identical behaviors. Along the same lines, institutions create a common set of incentives and constraints on behavior, which could encourage conformity (Bednar and Page, 2007; North, 1990; Young, 1998).

Often institutions or the environment create strong incentives for identical behavior. If everyone else in a community shakes hands upon greeting, drives on the left side of the road, and speaks English, an individual benefits from doing the same. In these instances, the incentives to take the same action as others are economic as opposed to psychological and are therefore considered as strategic coordination games and not instances of conformity. Not all conformity, though, can be seen as economically based. This holds true even in economic contexts. Young and Burke (2001), for example, show that rather than choose optimal contracts based on soil conditions, landowners and tenant farmers coordinate on common revenue sharing arrangements.

An incentive to conform also arises in uncertain situations, where individuals turn to others for behavioral clues (Liska, 1975; Tittle and Hill, 1967). In general, the more observable one's behavior is to others, the more likely one is to conform to the majority behavior and/or the standing social norm (Ajzen and Fishbein, 1969; Bowers, 1968; Liska, 1975; Warner and DeFleur 1969). If someone sees that everyone else has taken some action, she cannot help but draw inferences about the beliefs of others. This tendency can lead to herd behavior (Banerjee, 1992) and information cascades (Bikchandani et al., 1992).

The second fundamental force that we consider in our model relates to an individual's desire to be consistent. Moral principles may undergird consistency. For example, ideology or religion may provide umbrella beliefs or a set of values to guide behavior that are unlikely to change significantly over time.[8] The drive to consistency may also be based on convenience: consistent behavior across domains reduces cognitive load.[9] Internal consistency, whether motivated by principles or cognitive burden, provides the linkage across dimensions necessary for a recognizable cultural signature.

Like conformity, the desire for consistency is also well established empirically. Psychological research shows that personal uneasiness with cognitive dissonance creates within individuals a desire for consistency; people find it difficult to behave differently in every situation (Ajzen and Fishbein, 1980; Festinger, 1957; Ross and Ward, 1995; Van Overwalle and Jordens, 2002). Consistency extends to the political realm: Darmofal (2005) finds a citizen will tend to disagree with public policy experts' recommendations when they run counter to an individual's life experiences. Psychologists generally agree that individuals overcome cognitive dissonance by either restricting their behavior to be consistent with their attitudes or by changing their attitudes to match their inconsistent behavior (Aronson, 1999; Beauvois and Joule, 1996; Harmon-Jones and Harmon-Jones, 2002; McGuire, 1966; Singer, 1966).

Research in neuroscience complements psychological evidence for consistency. The physiology of the brain may enforce minimal levels of consistency for the simple reason that repeated behaviors create cognitive pathways which funnel future thought and action (Gazzaniga, 1999). This neurological evidence aligns with empirical studies that show when confronted with a novel situation, an individual often chooses a behavioral response that belongs to their existing repertoire (Cavalli-Sforza and Feldman, 1981; March, 1991).

Finally, at a more abstract level, consistency can be justified theoretically using the logic of cost-benefit analysis. Consistent behavior allows others to predict his/her next moves. Accurate predictions grease the wheels of

economic and political institutions. In fact, one broadly-accepted role of culture is to help coordinate on equilibria. Some equilibria may be more focal than others based on their relationship to the wider culture (Calvert and Johnson, 1998).

To summarize, empirical evidence shows that individuals tend toward both consistency and conformity. Note that we do not take consistency and conformity to be hard and fast rules but, rather, emphasize that they are two general forces that guide human behavior. Moment-to-moment, individuals can be as unpredictable in their behavior as they might be arbitrary in whom they imitate, but on the whole, they tend to be consistent and coordinate with others. Below, we present our model that captures these two forces that motivate human behavior. We show that these two simple pressures alone can generate the macro-observation of groups that are distinct from one another, that possess meaningful culture signatures, and that are characterized by persistent within-culture diversity.

2.1 General structure of the models

In the model, we assume N agents indexed by j. Each agent can be represented by a vector of M attributes that take one of A values. We write agent j as $a^j = a^j_1,\ a^j_2,\ldots\ a^j_M$, where each $a^j_i \in \{0,\ 1,\ldots\ A\}$. These attributes can represent behaviors, dispositions, customs, attire and so on. Crucial to our model will be that attribute values have meaning. If an agent's attributes take the same value, i.e. if $a^j_i = a^j_k$, then the agent will be consistent on those two attributes. The agents will alter their attributes over time following behavioral rules, which we describe below.[10] In what follows, we first describe pure internal consistency and pure social conformity models. We then combine the two models to create a consistent, conformity model. The single force models prove important to understand the process of cultural formation. To make that process as transparent as possible, in the mathematical analysis that follows we restrict attention to the case of binary attribute values, $a^j_i \in \{0,\ 1\}$.[11]

2.2 Force 1: Internal consistency

We first describe a model in which agents strive for consistency. We capture this formally as follows: We let $s(a^j)$ equal the number of times the most common attribute appears in agent j's vector of attributes. We can write a *consistency payoff function* as follows:

$$s(a^j) = max_{a \in A}\{|i|:\ a^j_i = a\}.$$

Given this payoff function, an optimizing agent would set all attributes to the same value. In our model, we implicitly assume that agents lack a holistic awareness of their inconsistencies. Instead, they alter their attribute values with the goal of increasing consistency. To model this process, in each period we randomly select an agent. The agent then applies the following *internal consistency rule*:

Internal consistency rule: *The agent chooses two random distinct attributes and changes the value of the first attribute to match the value of the second.*

Repeated application of this rule leads to consistency. To see why, assume attributes take value zero or one. Let x denote the number of an agent's attributes with value one. Recall that M attributes exist in total. Thus, if $x = 0$ or $x = M$, the agent is *consistent*. In game theory, these consistent states would be called *equilibria*, while in dynamical systems they would be called *absorbing states*. Given our focus on process, we adopt the latter terminology.

The internal consistency rule has the feature that the consistency payoff function is equally likely to increase as it is to decrease. Consider the special case where *exactly* one attribute has value one and all other attributes have value zero. The rule chooses two attributes with only the first changing its value. For the consistency to *increase* the first attribute chosen must be the single attribute with value one, in which case, the agent becomes fully consistent. This occurs with probability $\frac{1}{M}$. For consistency to *decrease*, the *second* attribute selected has to be the single attribute with value one. The probability that the first attribute has value zero equals $\frac{M-1}{M}$, and the probability that the second has value one equals $\frac{1}{M-1}$. The combined probability equals $\frac{1}{M}$, which equals the probability that consistency increased. We can now state the following lemma, whose proof relies on an extension of this logic.

Lemma 1 *Let x denote the number of M attributes whose values equal one. Applying the internal consistency rule, the probability that x increases or decreases by one equals:*

$$\frac{(M-x)x}{M(M-1)}$$

pf. For x to increase, the first attribute must be one of the $M - x$ attributes with value zero, and the second attribute must belong to one of the x attributes

with value one. These events occur with probabilities $\frac{M-x}{M}$ and $\frac{x}{M-1}$. The proof for the case in which x decreases follows the same logic.

Given this lemma, we can infer two features of the dynamics. First, because the number of ones is equally likely to increase or decrease, the process has no bias in either direction. Second, agents who are very inconsistent, i.e. who have nearly equal numbers of zeros and ones, relatively quickly become more consistent. In contrast, agents who are nearly consistent may take relatively long to become fully consistent. In technical terms, we can describe the process as an unbiased random walk in which the probability of movement slows near the two absorbing states.

2.3 Force 2: Coordination, social conformity

We next describe a pure conformity model. This model underpins many of the culture models. For the conformity model, we define the payoff to agent j, $f(a^j, a^{-j})$, as the percentage of other agents whose attributes match those of agent j averaged across all attributes:

$$f(a^j, a^{-j}) = \frac{\sum_{k \neq j} \sum_{i=1}^{M} \delta(a_i^j, a_i^k)}{NM}$$

where $\delta(a_i^j, a_i^k) = 1$ if and only if $a_i^j = a_i^k$. In other words, $\delta = 1$ if the agents' values agree on the attribute and zero if they do not. We refer to this as the *conformity payoff function.* In the model, in each period, we randomly choose a pair of agents. Note that this differs from the internal consistency rule in which we chose a pair of *attributes.* The first agent chosen applies the *social conformity rule:*

Social conformity rule: *The first paired agent randomly chooses an attribute and sets the value of that attribute equal to the value that the other agent assigns to that attribute.*

This rule converges to full conformity if the agents update asynchronously (Page, 1997). Like the internal consistency rule we established above, the social conformity rule also creates a random walk. The next lemma applies to a single attribute version ($M = 1$) of the model. The extension to the more general case is straightforward.

Lemma 2 *Let $M = 1$ and let y denote the number of N agents whose values on the attribute equal one. Applying the social conformity rule, the probability that y increases or decreases by one equals:*

$$\frac{(N-y)y}{N(N-1)}$$

The proof follows that of the consistency model because the processes are equivalent revealing a symmetry that can be made formal.

Observation: *The internal consistency model applied to N agents with M attributes is equivalent to the social conformity model applied to M agents with N attributes.*

In other words, a one dimensional pure conformity model is equivalent to a multidimensional pure consistency model. We now turn to the full model in which agents care about both consistency and conformity.

2.4 The forces combined: The consistent conformity model

In the combined model, we characterize the payoff function to an agent j, π_j, as a convex combination of the payoff functions of the first two models:

$$\pi_j(a^j, a^{-j}) = \alpha s(a^j) + (1 - \alpha) f(a^j, a^{-j})$$

where $\alpha \in [0, 1]$ denotes the relative weight on consistency. To model behavior we combine the previous two rules to create a single parameter family of rules $CC(p)$ where p denotes the probability that the agent applies the internal consistency rule. The parameter p may or may not equal α. As in the conformity model, in each period we randomly choose a pair of agents. The first agent in the pair applies the following behavioral rule:

Consistent conformity rule CC(p): *With probability p the activated agent applies the internal consistency rule and with probability $(1 - p)$ the activated agent applies the social conformity rule.*

Note that this construction makes the consistency and conformity models special cases of this model, where $CC(1)$ is the consistency model and $CC(0)$ is the conformity model. The equilibria given the payoff function described above are the same as the absorbing states given our behavioral rule: all agents choose the same value for each attribute.[12]

The next lemma characterizes the dynamics of the class of $CC(p)$ models. It states that the probability that an agent increases the number of attributes with value one depends on the number of attributes the agent sets to one (the consistency effect) and the number of other agents' attributes that equal one.

Lemma 3 *Assume a population of N agents with M binary attributes and an agent whose first x attributes take value one. Let S_i equal the number of*

other agents in the population who have value one on attribute i. The probability that x increases by one equals

$$p\frac{x(M-x)}{M(M-1)} + (1-p)\frac{1}{M}\sum_{i=x+1}^{M}\frac{S_i}{N-1},$$

and the probability that x decreases by one equals

$$p\frac{x(M-x)}{M(M-1)} + (1-p)\frac{1}{M}\sum_{i=1}^{M}\frac{N-1-S_i}{N-1}.$$

pf: The proof follows from the first two lemmas.

As in the two single force models, this dynamic bounces back and forth and slows near the absorbing states. One important difference between this dynamic and the others is that it has far fewer equilibria (absorbing states). To see why, recall that N is the number of agents, M is the number of attributes, and A is the number of values per attribute. In the consistency model there are A^M equilibria, in the conformity model, A^N equilibria, but in the combined CC(p) model, only A equilibria. The fact that the consistent conformity model has far fewer equilibria does not necessarily imply that those equilibria will take longer to locate, but when we expose the dynamics we'll see that this is in fact the case.

Before we present our analytic results, an example is instructive to demonstrate the tension between conformity and consistency. Suppose that two members of a society interact in three distinct contexts. In each context, a person can take a *fair* action, *F*, that equally splits resources, or take a *utilitarian* action, *U*, that produces a higher total payoff. Given these assumptions, we can describe an agent by a vector of length three consisting of *F*'s and *U*'s. Let's call these agents John and Jeremy, and assign them the following initial behavioral vectors:

John: (*F, F, U*)
Jeremy: (*F, U, U*).

Assume first that John and Jeremy apply the internal consistency rule. John may switch his third attribute so that his vector of attributes becomes (*F, F, F*). Jeremy, in contrast, may switch his first attribute so that his vector becomes (*U, U, U*). John and Jeremy would both achieve internal consistency and do so quickly.[13] Or, suppose John and Jeremy apply the social conformity rule. In this case, if John is activated first, and it is his second attribute that is selected, then John would switch his second attribute to *U* so that his vector becomes (*F, U, U*) The two agents have quickly reached conformity.

Now, assume that John and Jeremy apply both the internal consistency and the social conformity rules. John might first switch to (*F, F, F*). He may then meet Jeremy and switch to (*F, U, F*). However, he may then realize that he is being inconsistent and switch back to (*F, F, F*). Jeremy, meanwhile, may switch to (*U, U, U*) and then, aiming to conform, switch back to (*F, U, U*). Eventually, both John and Jeremy will be consistent and conform with one another but it would take much longer. Because conformity and consistency can pull in different directions, the time required to attain an equilibrium can be greatly increased.

3 Analytic results

We now establish analytical claims on time to convergence. In this section we demonstrate how the process of cultural signature creation may be lengthy by considering a model without any errors, an assumption we relax in the next subsection. We compare the time to convergence in the three models – pure consistency, pure conformity, and the combined CC(p) model. We find that the combined model, necessary for cultural coherence, takes far longer to converge. We also vary the weight of the two pure models within the CC(p) model, finding that the greater the imbalance, the longer the convergence takes. We then introduce errors and show that the stochastic equilibrium of the combined model produces much more persistence of diversity than either of the two individual force models.

Often, time to convergence can serve as a proxy for complexity of underlying dynamics; processes that take longer to converge have more interesting dynamics.[14] Further, more complex dynamic processes are more likely to amplify small errors and create persistent heterogeneity. In our model, we find this to be true. In fact, time to convergence proves a *perfect* correlate to error amplification.[15]

We consider the simplest interesting case: two agents, two attributes, and two values per attribute ($A = 2$, $N = 2$, $M = 2$). This case proves sufficient to demonstrate two main results: (1) the consistent conformity model takes longer to converge than either of the two single force models, and (2) its equilibrium in the model with errors has greater dispersion.[16]

With two agents and two binary attributes, the systems can be in any one of sixteen states which can be sorted into five categories: the two agents can be in conformity and internally consistent (C&C), consistent but not conforming (CON), conforming but not consistent (CRD), one agent can be consistent but the other not, what we call off by one (OBO), or both can be inconsistent and lack conformity (NOT). Using the letters *a* and *b* to denote distinct attribute values, in Table 1 we define each category and its probabilities, given random assignment of attributes.

Table 1. States of the system

State	Agents	Probability
Conformed & Consistent (C&C)	(a,a) (a,a)	$\frac{1}{8}$
Consistent Not Conformed (CON)	(a,a) (b,b)	$\frac{1}{8}$
Conformed Not Consistent (CRD)	(a,b) (a,b)	$\frac{1}{8}$
Off By One (OBO)	(a,b) (a,a)	$\frac{1}{2}$
Not Conformed Not Consistent (NOT)	(a,b) (b,a)	$\frac{1}{8}$
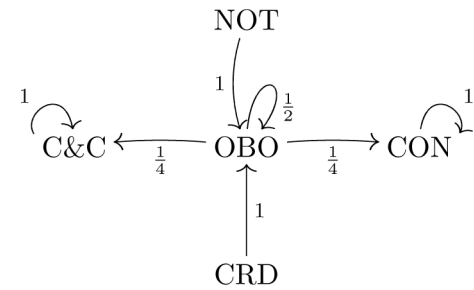
Figure 2. The dynamics of the internal consistency rule

Consistency model

In the consistency model, the equilibria consist of configurations in the sets *C&C* and *CON*. We can calculate the probability that any one of the other states moves to those states. For example, if the initial state is in *OBO*, then applying the internal consistency rule, the probability of staying in *OBO* equals one half, and the probability of moving to *C&C* or to *CON* equals one fourth. Similarly, if the initial state is *NOT* or *CRD*, then it moves into *OBO* with probability one. We display this information diagrammatically in Figure 2.

Notice how the flow diagram shows the potential for bouncing back and forth in the category *OBO* before eaching an absorbing state. That oscillation is limited in expected duration, as we now show in a lemma.

Lemma 4 *With two agents and two attributes, the expected time to equilibrium for the Internal Consistency Rule is* $1\frac{3}{4}$ *interactions.*[17]

pf: Let T_s denote the time (or expected time) to get to equilbrium from a given state. First, note that $T_{CON} = T_{C\&C} = 0$, since *C&C* and *CON* are absorbing states. Second note that the time to reach an absorbing state from a state in *CRD* or *NOT* equals one plus the time it takes to reach an absorbing state from *OBO*.

$$T_{CRD} = T_{NOT} = 1 + T_{OBO}$$

We calculate the expected time to reach an absorbing state from *OBO* as follows: With probability one half, the process takes only one time period. The other half of the time, the process remains in *OBO*, which means the time to an absorbing state equals one plus the time to an absorbing state from *OBO*. We can write this as follows:

$$T_{OBO} = \frac{1}{2}(1) + \frac{1}{2}(1 + T_{OBO}) = 1 + \frac{1}{2}T_{OBO}$$

Solving for T_{OBO} yields that $T_{OBO} = 2$. Therefore $T_{CRD} = T_{NOT} = 3$, so applying the internal consistency rule, the expected time to attain an absorbing state, T^{ICR}, equals $T^{ICR} = \frac{1}{8}(0) + \frac{1}{8}(0) + \frac{1}{8}(3) + \frac{1}{8}(3) + \frac{1}{2}(2) = 1\frac{3}{4}$.

Conformity model

We next construct a similar diagram for the dynamics created by the social conformity rule (see Figure 3). Notice that Figure 3 is identical to Figure 2 except that the states *CRD* and *CON* have changed places. In light of that symmetry, it follows that the expected time to an absorbing state in this model is also $1\frac{3}{4}$ interactions.

Lemma 5 *With two agents and two attributes, the expected time to equilibrium for the Social Conformity Rule is* $1\frac{3}{4}$ *interactions.*

pf: follows from above.

CC(p) Model

Next, we consider the combined coordination/consistency model, or *CC* (*p*). In the diagram below, we show the case where $p = \frac{1}{2}$. The diagram for this model, Figure 4, combines the diagrams for the previous two models. Notice that the only absorbing state (equilibrium) is *C&C*. Using Figure 4, we can state the following claim.
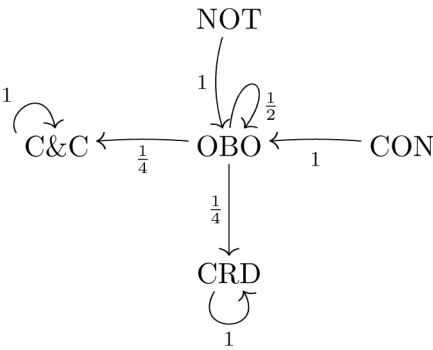
Figure 3. The dynamics of the social conformity rule
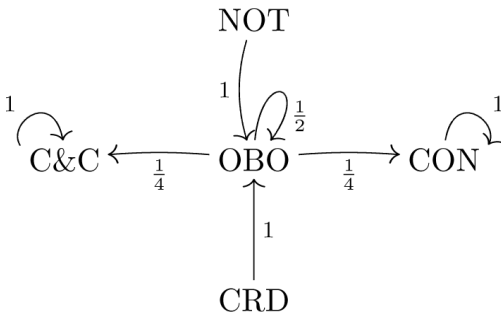
Figure 4. The dynamics of the combined rule where $p = \frac{1}{2}$

Lemma 6 *With two agents and two attributes, the expected time to equilibrium for the CC(p) Rule is* $1\frac{7}{8} + \frac{1}{p(1-p)}$.

For proof, see appendix.

Our first claim follows from the previous three lemmas.

Claim 1 *With two agents and two attributes, the expected time to equilibrium in the combined model is strictly greater than either the social conformity or internal consistency models.*

pf: Follows directly from Lemmas 4, 5, and 6.

We can also use Lemma 6 to establish a second substantive claim and corollary that compare the time to convergence as we vary the probability that the consistency rule is invoked:

Claim 2 *In the CC(p) model, the time to convergence is minimized at* $p = \frac{1}{2}$.

Corollary 1 *For* $p \in (0,1)$, *the time to convergence increases as* $p \to 0$ *or* $p \to 1$.

Proof: Straightforward from Lemma 6.

Given that we have equal numbers of agents and attributes, if the behavioral rule includes both conformity and consistency, the time to convergence is minimized when the two forces are equally weighted. We can compare the expected time to equilibrium in the three models – including the full family of submodels in $CC(p)$ – graphically. Recall that the internal consistency model and the social conformity model are special cases of the combined model, where $p = 1$ or $p = 0$, respectively. Figure 5 shows the expected time to equilibrium as a function of the probability of applying the consistency rule. Notice that the expected time to equilibrium is far shorter in the conformity model and the consistency model than in the $CC(p)$ model, as established by Claim 1. For comparison, the time to convergence at $p = \frac{1}{2}$ is $5\frac{7}{8}$ interactions; this value is more than three times the time to convergence in the other two cases. Also note the non-linearity of the time to convergence – what we have taken to calling the 'smile' curve – due to the increasing time produced by an unbalanced weighting of the two rules.
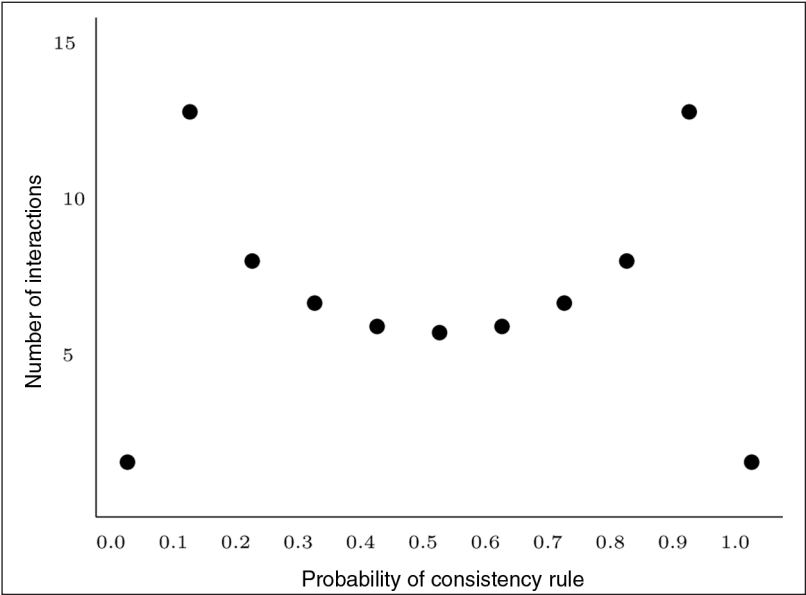
Figure 5. Expected time to equilibrium: two person model

The three flow diagrams reveal the two reasons why the consistent conformity model takes longer to converge than either the conformity model or the consistency model. First, as already proved, the consistent conformity model has fewer absorbing states. Whereas Figures 2 and 3 each have two categories of absorbing states, Figure 4 has a single category of absorbing states. Second, the individual processes of the pure conformity model and the pure consistency model head directly to their respective absorbing states. The only possible delay that can occur in either of those single processes is the system remaining in state *OBO*. In contrast, the consistent conformity model can move from *CRD* to *OBO* to *CON* and back to *OBO*, thus considerably slowing convergence.

Slow convergence can result from either a slow dynamic or complex interactions. In the case of the CC(p) rule, it is the latter cause. When we add errors to the dynamic, these complex interactions can enable those errors to accumulate resulting in greater variation in the limiting distribution. In other words, the more complex dynamics of the CC(p) model may result in greater intra-cultural heterogeneity. We make that connection more formal in the next section.

3.1 The persistence of diversity

The inclusion of errors is a standard assumption in learning and conformity models. These models still attain equilibria, but the equilibria are no longer static. Instead, the system reaches a limiting equilibrium distribution over states (Young, 1998). For example, with error, the conformity model might spend 95% of the time in the conformity state and 5% of the time in the category *OBO*. That would be an equilibrium distribution but not a static equilibrium.

To capture errors, we assume that with some small positive probability, ϵ, an agent randomly changes an attribute's value rather than applying its behavioral rule. We focus on how the forces for coordination and consistency singly and jointly magnify these errors. One might expect that adding noise at a level ϵ would create an equilibrium distribution in which approximately ϵ of the agents are out of equilibrium. And in the consistency model and the conformity model, we find something close to that. In the consistent conformity model, however, the behavioral rule magnifies the noise term.

Consistency model

We begin with the pure consistency model. It suffices to consider a single agent, which allows us to reduce our five states to three. We can let *CNS*

denote the union of the states *CON* and *C&C*. These represent the states where the agents are consistent. We can then combine the *NOT* and *CRD* into the state *NCN* . In this state, neither agent is consistent. This gives a Markov Process defined over three states *CNS*, *NCN*, and *OBO*. We can write the Markov transition matrix as follows:

		$T+1$		
		CNS	*OBO*	*NCN*
	CNS	$1-\epsilon$	ϵ	0
T	*OBO*	$\frac{1}{2}$	$\frac{1-\epsilon}{2}$	$\frac{\epsilon}{2}$
	NCN	0	1	0

This gives three equations that characterize the dynamic equilibrium:

$P_{CNS} = (1-\epsilon)P_{CNS} + \frac{1}{2}P_{OBO}$, $P_{OBO} = \epsilon P_{CNS} + \frac{1-\epsilon}{2}P_{OBO} + P_{NCN}$, and $P_{NCN} = \frac{\epsilon}{2}P_{OBO}$. Solving these equations gives

$$P_{CNS} = \frac{1}{1 + 2\epsilon + \epsilon^2}, P_{OBO} = \frac{2\epsilon}{1 + 2\epsilon + \epsilon^2}, P_{NCN} = \frac{\epsilon^2}{1 + 2\epsilon + \epsilon^2}$$

These last three equations characterize the dynamic equilibrium. The proportion of agents in the consistent state, *CNS*, equals approximately $\frac{1}{1 + 2\epsilon}$, which is approximately $1+2\epsilon$. In other words, ϵ error translates into 2ϵ of the population on average not in the error-free absorbing state.

Conformity model

We can perform a similar analysis for the conformity model. Let *CDC* equal the union of the two states in which the two agents have conformed, *CRD* and *C&C*, and let *NCD* equal the union of the states in which they have not, *NOT* and *CON*. We can write the Markov transition matrix as follows:

		$T+1$		
		CDC	*OBO*	*NCD*
	CDC	$1-\epsilon$	ϵ	0
T	*OBO*	$\frac{1}{2}$	$\frac{1-\epsilon}{2}$	$\frac{\epsilon}{2}$
	NCD	0	1	0

This matrix is identical to the one for the consistency model up to a relabeling of the states. Therefore, the equilibrium equals

$$P_{CDC} = \frac{1}{1 + 2\epsilon + \epsilon^2}, P_{OBO} = \frac{2\epsilon}{1 + 2\epsilon + \epsilon^2}, P_{NCD} = \frac{\epsilon^2}{1 + 2\epsilon + \epsilon^2}$$

These equations can be interpreted the same as the previous ones: introducing an error of size ϵ produces an equilibrium that has on average 2ϵ of the agents not in the error-free model's absorbing state.

CC(p) model

For the consistent conformity model, we can write the Markov transition matrix as follows:

|		T+1				
		C&C	OBO	CRD	CON	NOT
	C&C	$1 - \epsilon$	ϵ	0	0	0
	OBO	$\frac{1}{4}$	$\frac{1-\epsilon}{2}$	$\frac{p+\epsilon-\epsilon p}{4}$	$\frac{1-p-\epsilon p}{4}$	$\frac{\epsilon}{4}$
T	CRD	0	$1-p+\epsilon p$	$p-\epsilon p$	0	0
	CON	0	$p+\epsilon+\epsilon p$	0	$1-p-\epsilon+\epsilon p$	0
	NOT	0	1	0	0	0

The equilibrium of this process can be written as follows:

$$P_{C\&C} = \frac{1}{1 + 4\epsilon + \epsilon^2 + \alpha\epsilon + \alpha^{-1}\epsilon}, P_{OBO} = \frac{4\epsilon}{1 + 4\epsilon + \epsilon^2 + \alpha\epsilon + \alpha^{-1}\epsilon},$$

$$P_{CRD} = \frac{\alpha\epsilon}{1 + 4\epsilon + \epsilon^2 + \alpha\epsilon + \alpha^{-1}\epsilon}$$

$$P_{CON} = \frac{\alpha^{-1}\epsilon}{1 + 4\epsilon + \epsilon^2 + \alpha\epsilon + \alpha^{-1}\epsilon}, P_{NOT} = \frac{\epsilon^2}{1 + 4\epsilon + \epsilon^2 + \alpha\epsilon + \alpha^{-1}\epsilon}$$

Where $\alpha = \frac{(p + \epsilon - \epsilon p)}{(1 - p + \epsilon p)}$ which equals the ratio of the probability of moving from *OBO* to *CON* to the probability of moving from *OBO* to *CRD*. By inspection, $(\alpha + \alpha^{-1})$ must be greater than or equal to two.

Note first, and most importantly, that the proportion of time in the absorbing state equals approximately, $\dfrac{1}{1 + (4 + \alpha + \alpha^{-1})\epsilon}$. For small ϵ this is approximately $1 - (4 + \alpha + \alpha^{-1})\epsilon$. Thus, an error of ϵ implies that, on average, *at least* 6ϵ of the population is not in the error-free absorbing state. Not only does diversity persist, but at much higher rates than the level of error. Note second that the higher α, the more time the system will spend in *CON*. The lower α, the more time that the system will spend in *CRD*. Finally, note that setting $p = \dfrac{1}{2}$ maximizes the time spent in the consistent conformity state (*C&C*). Figure 6 shows the percentage of the time the system spends outside of state *C&C* as a function of p for a given error level ϵ. If we let p go to 0 then α converges to ϵ and the system spends half of the time outside of the state *C&C*. Similarly, if we let p go to 1 then α converges to $\dfrac{1}{\epsilon}$, and the system again spends half of the time outside of the state *C&C*. Thus, even for very small errors, if one dynamic outweighs the other, the system can spend almost half of its time out of equilibrium.

Except for the units on the *y*-axis, this figure matches Figure 6 exactly. The equivalence, modulo a rescaling, of the time to equilibrium and the
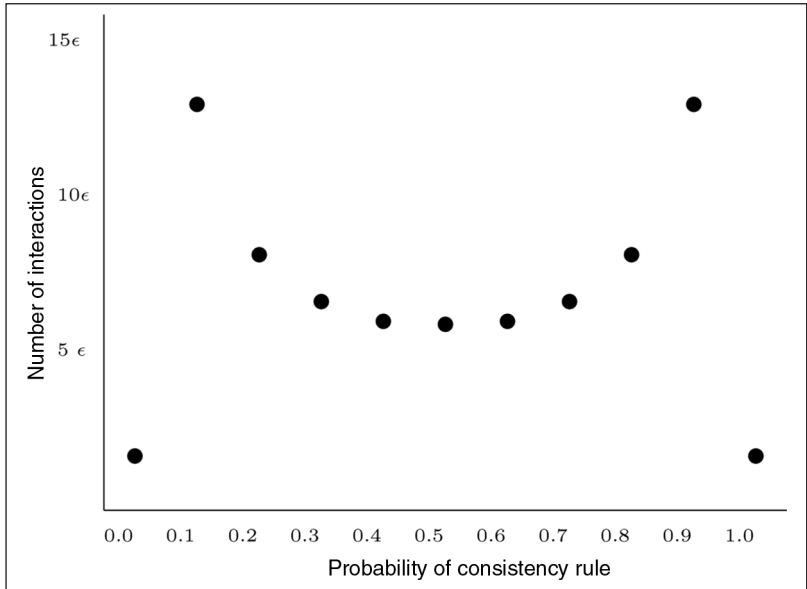

Figure 6. Distance to conformed and consistent equilibrium: error model

distance to the perfectly conformed and consistent equilibrium is an artifact of our model. But the correlation between the two generally hints at an important insight: the longer the time to equilibrium, the more complex the dynamics. The more complex the dynamics, the larger the potential effects of error.

We can summarize these results in the following claim:

Claim 3 *With two agents and two attributes, the average proportion of time not spent in an absorbing state in the social conformity or internal consistency models approximately equals 2ε. In the combined model, the average proportion of time not spent in an absorbing state is at least 6ε and can approach one-half for any ε.*

This claim establishes that small amounts of error produce substantial heterogeneity in the combined model. It also establishes that the model is capable of producing agents who are somewhat consistent and somewhat
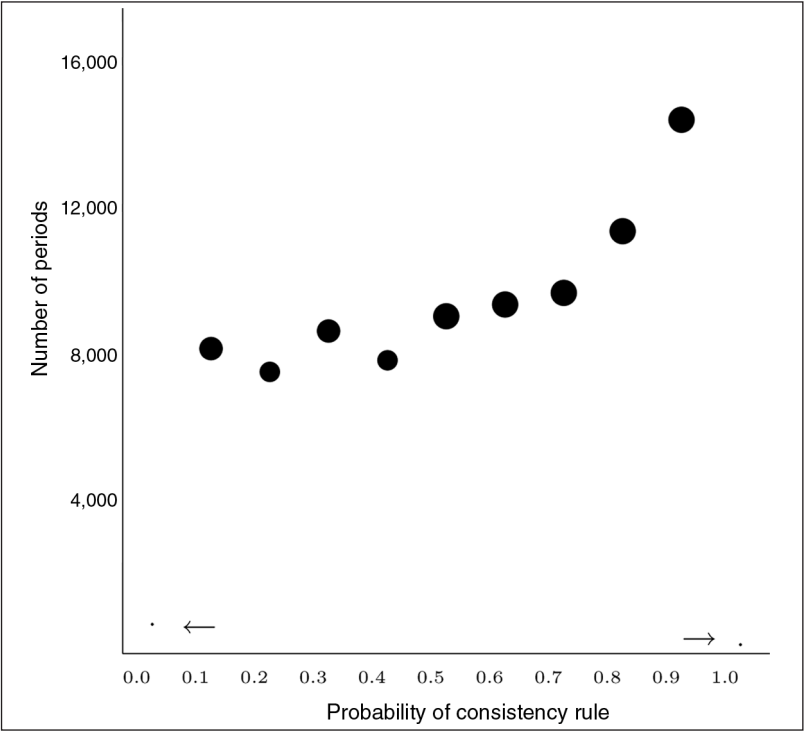
Figure 7. Time to convergence in number of periods

coordinated. A natural question to ask is whether this result scales: Do similar findings hold with larger number of agents and attributes? We provide a statistical answer to that question using numerical experiments.

4 Numerical experiments

We now run numerical experiments to test the robustness of our results to an increase in the number of agents and attributes. We varied the number of agents from two to one thousand, the number of attributes from two to ten, and the number of values per attribute from two to six.[18]

We present two sets of computational experiments. In the first set, we measure the time to convergence in the error-free model. In the second set, we measure levels of consistency and conformity in the models with errors. Figure 7 shows the time to convergence as a function of p for a model with one hundred agents, ten attributes, and six values per attribute. The results are averages of over fifty trials. All of the differences are statistically significant. The arrows point to the values for $p = 0$ and $p = 1$, which are otherwise easy to overlook.

Our theoretical results suggested that the time to convergence should increase as p approaches zero and one. Here, we only see that phenomenon as p approaches one because the number of agents is far larger than the number of attributes. The probability of applying the consistency rule must be very small before we would expect to see the time to convergence to increase given the greater need for conformity. All of the models converge, but the magnitude of the time to convergence differs substantially. In the conformity model and the consistency model, the system converges in a few hundred periods. The consistent conformity model can take more than fifteen thousand periods to converge.

In the second set of experiments we test to see whether errors have a much larger effect in the consistent conformity model. To make this comparison we need some measures of consistency and conformity. In constructing these measures, we refer back to notation we used in constructing payoff functions. Recall that $s(a^j)$ equals the number of times the most common attribute appears in agent j's vector of attributes. We define *pconsistent* as follows:

$$pconsistent = \frac{\sum_{j=1}^{M} s(a^j)}{AM}$$

Table 2. Percentage of conformity and consistency (models with error)

Noise	p = 0.0				p = 0.5				p = 1.0			
	pconformity		pconsistent		pconformity		pconsistent		pconformity		pconsistent	
	mean	stdev	mean	stdev	mean	stdev	mean	stdev	mean	stdev	mean	stdev
0	1	0	0.360	0.082	1	0	1	0	0.200	0.016	1	0
0.005	0.736	0.064	0.373	0.044	0.354	0.081	0.556	0.067	0.199	0.012	0.970	0.009
0.01	0.585	0.052	0.376	0.030	0.299	0.037	0.510	0.033	0.200	0.012	0.946	0.012
0.02	0.482	0.044	0.376	0.023	0.269	0.017	0.483	0.017	0.201	0.012	0.904	0.017

By construction, *pconsistent* takes on values in the closed interval between zero and one, with perfectly consistent agents taking the value one. We similarly define pconformity to be the average of the conformity payoff functions.

$$\text{pconformity} = \frac{\sum_{j=1}^{M} f(a^j, a^{-j})}{M}$$

Thus, if the entire population has conformed, then the value of *pcoordinate* equals one. Table 2 shows the average percent values and standard deviations of inter-agent value difference (pconformity) and intra-agent value difference (pconsistent) for each of the three models. The data cover the last 1000 interactions of 100 runs with 100 agents, 10 features, 5 values per feature and a total run time of 5,000,000 interactions per run. Notice that in alignment with theoretical predictions, with no errors, the $CC(\frac{1}{2})$ model converges to a consistent and coordinated state as we expect, and the two pure models (the equivalent of CC(0) and CC(1)) converge unproblematically, as predicted by the theory.

Most importantly, for the $CC(\frac{1}{2})$ model, the introduction of even the tiniest bit of noise (0.005) leads to substantial heterogeneity both between agents (only 35.4% conform) and within agents (55.6% are consistent), displaying far more diversity than in the other two models (73.6% conform and 97% are consistent). A little noise has a much larger effect when both forces operate.[19] These computational experiments show that the insight generated in the starker mathematical model – that the effect of noise when both forces are in play greatly exceeds the sum of the individual effects – becomes amplified in larger systems. Thus, even small amounts of error may produce substantial heterogeneity in a society of people who wish to conform and be consistent.

While our model offers an explanation for the emergence of cultural signatures and internal cultural diversity, it remains consistent with models based upon Axelrod (1997) that support differences between cultures. Two runs of any coordination game models (including ours) are very unlikely to produce the same outcomes. Thus, if we treat one run of the model as the United States, and a second run of the model as Germany, then we will naturally see two different 'cultures' emerge, each with its own signature. And because our model includes conformity, we also see that its distinct signatures are meaningful and consistent with the broader fabric of behaviors in that society. Also because of consistency, we see the persistence of diversity within each culture. As different runs of the model produce different outcomes, the model naturally provides an explanation for inter-cultural

diversity – differences in initial conditions and different paths lead to diverse outcomes. Consistent with Axelrod (1997), even when these cultures interact, group distinctions persist.

5 Conclusion

In this paper, we construct a model that produces four hallmarks of culture: coordinated behavior, coherent cultural signatures, substantial within culture diversity, and cross cultural differences. That model combines a social drive to coordinate with an individual desire for internal consistency. In our analysis, we explore the dynamics of the model and show how the combined forces for conformity and consistency result in greater complexity and longer time to convergence than found in models of culture that consider only coordination.

We also find that varying the weights of the forces slows convergence and increases diversity when the model includes error. This finding may help us to understand why some cultures exhibit more diversity than others: if pressures to conform or be consistent are higher or lower, we'll see different levels of diversity. For example, in a society in which the relative tendency to conform is high relative to the tendency to be consistent, people may be less consistent but more similar. Thus, whether one culture appears more or less heterogeneous depends on the type of questions asked in a survey. If survey questions ask about an existing behavior, we'd expect a higher conforming society to appear less heterogeneous. However, if the questions are hypothetical, the lack of consistency may give respondents a variety of possible behaviors to apply in the novel context. Depending on the questions asked, a less individualistic society, like Japan, could appear more heterogeneous than a highly individualistic society like the United States.

Relatedly, the presence of more or fewer 'errors' (such as propensities towards missteps, misinformation, or confusion) in different systems would respectively promote or inhibit diversity. We might then speculate that informational systems provide a crude proxy for the transmission error of cultural traits. Closer relations between individuals would push in the opposite direction.

As mentioned in the introduction, in our model, we ascribe meaning to the attribute values. A worthwhile extension of this model would be to allow for the meaning of attribute values to be endogenous, to be socially constructed. To accomplish that we could differentiate between the set of *actions* on attributes and the *values* associated with those actions. Meaning making would entail mapping actions into values. That meaning making

must be constrained in some way. If the maps could be arbitrary, then outsiders would not be capable of external validation of consistent cultural signatures. Thus, an extension must include enough flexibility in meaning making to allow for substantial reinterpretations of actions but at the same time, not so much flexibility that cultural signatures fail to exist. We hasten to add that such a model would, among other results, probably produce greater intra-cultural heterogeneity owing to the coordination on meaning that must also occur.

Our model also has implications for a range of organizational forms. Within corporations, for example, people face incentives to conform as well as to be consistent, though for reasons that differ slightly from those we described above. Relatedly, members of a political party also desire conformity and consistency, and these two desires may result in the analogous effects: differences within and between parties as well as coherent party ideologies. In a party version of the model, attribute values would represent participants' ideal points in policy or preference space. The internal consistency rule would capture the individual desire for a consistent ideology, and the social conformity rule would capture the collective desire for party cohesiveness. The simplicity of our model means it could be reasonably extended in a number of ways to be sensitive to any particular constraints in different contexts in which it is relevant.

One implication of our results is that consistent cohesiveness cannot emerge without top down encouragement or even enforcement.[20] Within any organization or collection of people, be it an interest group, a community organization, or an academic department, these two forces probably operate. Absent strong central control, diversity should reign.[21] This finding agrees with what we see in the real world: few (if any) groups converge to a state of consistent conformity, but meaningful cultural signatures do emerge.

Intra-group heterogeneity undoubtedly influences political and economic outcomes. Most empirical studies of heterogeneity and its effects focus on ethnic or religious differences. Here, we're considering variations within such groups. So in our case, the lack of convergence, be it in a society, a political party, or an organization, may, on balance, be a good thing. It may promote innovation in the form of cultural evolution. It may even be interpreted as a form of dissent as individuals either conform or stick to their own lights. This tension between conformity and consistency maps to related tensions between 'exploiters versus explorers', 'conformers versus nonconformers', and 'scroungers versus producers' and may produce stability through variation (Axelrod and Cohen, 2000; Boyd and Richerson, 1985; Campbell, 1965; Kameda and Nakanishi, 2002; March, 1991; Nisbett and Ross, 1980; Roberts, 1964; Rogers, 1995; Tindall, 1976; Weick, 1969).

Diverse societies may also better produce knowledge and be more adept at problem solving and prediction (Page, 2007). Overall, diverse societies may be more robust, as they have the potential to adapt to new and changing circumstances (Bednar, 2009). In contrast, societies that lack intra-cultural diversity may be prone to collapse (Diamond, 2005). Thus, the persistence of diversity in the face of two homogenizing forces may prove as serendipitous as it is paradoxical.

Appendix

Claim 4 *The expected time to convergence for the consistency model with binary values and M attributes for a random starting point is of order M^2 periods.*

pf: (Courtesy of Len Sander) Let x denote the number of attributes with value 1. Let T_x be the time to convergence if at location x. Let m_x be the probability of increasing or decreasing the number of attributes with value 1. By the previous claim, these probabilities are equal. After one time period, the expected time has to be one period less. Therefore, we have the following equation:

$$T_x - 1 = m_x T_{x+1} + m_x T_{x-1} + (1 - 2m_x) T_x$$

This reduces to

$$-1 = m_x [(T_{x+1} - T_x) - (T_x - T_{x-1})]$$

Recall from Claim 1 that $m_x = \dfrac{(M-x)x}{M(M-1)}$. For large M we can approximate this as $m_x = \dfrac{(M-x)x}{M^2}$. Let $p(x) = \dfrac{x}{M}$, so that $m_x = p(x)[1 - p(x)]$. We then can rewrite $T_{x+1} - T_x$ as

$$\frac{1}{M} \cdot \frac{(T(p(x+1)) - T(p(x)))}{\dfrac{1}{M}}$$

For large M, this converges to $\dfrac{\partial T(p(x))}{\partial p}$. It follows that we can write the following approximation:

$$(T_{x+1} - T_x) - (T_x - T_{x-1}) \sim \frac{1}{M}\left[\frac{\partial T(p(x))}{\partial p} - \frac{\partial T(p(x-1))}{\partial p} \right]$$

Which in turn we can approximate as

$$\frac{1}{M^2}\frac{\partial^2 T(p(x))}{\partial p}$$

We can therefore approximate our initial difference equation as

$$-1 = p(x)[1 - p(x)]\frac{1}{M^2}\frac{\partial^2 T(p(x))}{\partial p}$$

Rearranging terms and simplifying notation gives

$$\frac{\partial^2 T(p)}{\partial p^2} = -\frac{M^2}{p(1-p)}$$

We also have that $T(0) = T(1) = 0$. The solution to this differential equation is

$$T(p) = M^2\left[p\log\left(\frac{1}{p}\right) + (1-p)\log\left(\frac{1}{1-p}\right)\right]$$

which completes the proof.

We can state a similar result for the conformity model.

Corollary 2 *For the conformity model with binary values and N agents, the expected time to convergence for a random starting point is of order N^2 periods.*

pf: follows from our earlier observation of equivalence and the previous claim.

Proof of Lemma 6:

We can write the following equations.

$$T_{C\&C} = 0$$

$$T_{OBO} = 1 + \frac{1}{4}T_{C\&C} + \frac{1}{2}T_{OBO} + \frac{p}{4}T_{CON} + \frac{(1-p)}{4}T_{CRD}$$

$$T_{CON} = 1 + (1-p)T_{OBO} + pT_{CON}$$

$$T_{CRD} = 1 + pT_{OBO} + (1-p)T_{CRD}$$

$$T_{NOT} = 1 + T_{OBO}$$

By substitution, these equations imply that

$$T_{CON} = \frac{1}{1-p} + T_{OBO}, \text{ and } T_{CRD} = \frac{1}{p} + T_{OBO}$$

These in turn imply that

$$T_{OBO} = 1 + \frac{1}{2}T_{OBO} + \frac{p}{4(1-p)} + \frac{(1-p)}{4p} + \frac{1}{4}T_{OBO}$$

This reduces to

$$T_{OBO} = 4 + \frac{(1 - 2p + 2p^2)}{p(1-p)}$$

Substituting back into the other equations gives

$$T_{CON} = 4 + \frac{(1 - p + 2p^2)}{p(1-p)}, T_{CRD} = 4 + \frac{(2 - 3p + 2p^2)}{p(1-p)},$$

$$T_{NOT} = 5 + \frac{(1 - 2p + 2p^2)}{p(1-p)}$$

Therefore the average time to convergence equals

$$\frac{1}{2}\left(4 + \frac{(1 - 2p + 2p^2)}{p(1-p)}\right) + \frac{1}{8}\left(\begin{array}{c} 4 + \dfrac{(1 - p + 2p^2)}{p(1-p)} + 4 + \dfrac{(2 - 3p + 2p^2)}{p(1-p)} \\ +5 + \dfrac{(1 - 2p + 2p^2)}{p(1-p)} \end{array}\right)$$

Which reduces to

$$1\frac{7}{8} + \frac{1}{p(1-p)}$$

For the special case $p = \frac{1}{2}$, these equations become

$$T_{OBO} = 1 + \frac{1}{4}T_{C\&C} + \frac{1}{2}T_{OBO} + \frac{1}{8}T_{CON} + \frac{1}{8}T_{CRD}$$

$$T_{CON} = 1 + \frac{1}{2}T_{OBO} + \frac{1}{2}T_{CON}$$

$$T_{CRD} = 1 + \frac{1}{2}T_{OBO} + \frac{1}{2}T_{CRD}$$

$$T_{NOT} = 1 + T_{OBO}$$

By substitution, these equations imply that $T_{CON} = T_{CRD} = 2 + T_{OBO}$. Which in turn implies that $T_{OBO} = 1 + \frac{1}{2}T_{OBO} + \frac{1}{2} + \frac{1}{4}T_{OBO}$. This is an equation in a single variable, T_{OBO}. Solving this equation gives $T_{OBO} = 6$ Substituting back into the other equations gives $T_{CON} = T_{CRD} = 8$ and $T_{NOT} = 7$. Therefore the average time to convergence equals $\frac{1}{2}(6) + \frac{1}{8}(8 + 8 + 7) = 5\frac{7}{8}$.

Acknowledgements

Notes
1. In addition to interaction and influence dynamics, many of these models show that variations in network configurations of agents and/or network dynamics (for example, completely cutting off relations with a neighbor with whom one shares no common traits) can also influence the extent to which global polarization and local convergence take place.
2. Interpretative cultural theorists such as Sewell (1999) and Wedeen (2002) refer to a related construct of coherence. Their formulation allows for substantial social construction of meaning.
3. In coordination models, the result that people break into subcultures is a scaled down version of a global/local model of between-culture heterogeneity. In contrast, when we write of heterogeneity and diversity in this paper, we refer to differences between individuals in a population.
4. These findings are consistent with those of other conformity models, including those that allow for preferential interaction (Axelrod, 1997; Barth, 1969; Hannan, 1979; Homans, 1950; Macy et al., 2003; Rogers, 1983; and Simmel, 1955).
5. On most variables he finds significant variation between country means. On cross cultural differences in life satisfaction over 64 countries, the United States life satisfaction mean is 7.7, based on a ten-point scale; across all 64 societies the means range from as low as 3.7 to as high as 8.2.
6. Bandura (1977) demonstrated that children imitate behavior they view on TV and Huesmann (1988, 1998) showed that copied behaviors become encoded into children's behavioral schemas.
7. McPherson et al. (2001) review the scores of empirical studies. See also Schelling's (1971) discussion of preferences and racial segregation.
8. Note that these umbrella beliefs do not imply perfect consistency. People are prone to idiosyncratic changes in behavior (or what is 'consistent' for one person might not be consistent for another), much as they might idiosyncratically copy a neighbor or not. The inclusion of errors in the second version of the model helps capture this.

 9. See Bednar and Page (2007) for a discussion of cognitive load and culture.
10. Thus, these attributes are not *sacred* in the sense of Ginges et al. (2007).
11. The extension to non-binary attributes is notationally burdensome but straightforward.
12. Note that other behavioral rules, in particular a myopic best response adjustment process, in which an agent only switches an attribute's value if it leads to a higher payoff, can produce inefficient equilibria as well (Kuran and Sandholm, 2008).
13. Note that John could also change to (*F, U, U*) or (*U, F, U*) given the internal consistency rule, but at some point, he would have all three of his attributes taking the same value.
14. In studies of large systems, such as entire societies, equilibria may not be attained in reasonable time. Thus, time to convergence merits attention on its own (Page, 2008).
15. The class of problems that we consider here has analogs in physics and computer science. Scholars in those fields have analyzed time to convergence as the number of attributes or agents grows very large. Using techniques developed by Bouchaud et al. (1990), it can be shown that the time to convergence in the consistency model is of order M^2, and time to convergence in the conformity model with one attribute is of order N^2 (see appendix). The time to convergence for the consistent conformity model can be shown to increase in order $N^2 M^2$ for $P = 1/2$ (Page et al., 2007).
16. A similar analysis for the three binary attribute, two agent model demonstrates the same qualitative results with the quantitative findings exaggerated.
17. Time is measured by the number of interactions (an interaction is one application of a rule) with each interaction taking one time step. Hence, time is really a measure of the iterations of the model irrespective of the computational complexity of the iteration.
18. As a robustness check, we wrote two separate programs, one in *C* and one in Repast (a java-based modeling toolkit). The models produced indistinguishable results. All results presented come from the model written in *C*.
19. Comparing results for cases with noise = 0.005, the p-value for a test of the difference of means for conformity for the p = 0.0 and p = 0.5 models is 2.23×10^{-304} and the p-value for a test of the difference of means for consistency for the p = 0.5 and p = 1.0 models is 1.61×10^{-831}.
20. As presented, the model considers random mixing and no central transmission of desired attributes.
21. Note that that control would need to be very strong, as our model shows a preponderance of incentives to conform typically slows convergence.

References

Ajzen I and Fishbein M (1980) *Understanding Attitudes and Predicting Social Behavior*. Englewood Cliffs, NJ: Prentice-Hall.

Aronson E (1999) Dissonance, hypocrisy, and the self concept. In: Harmon-Jones E and Mills J (eds) *Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology*. Washington, DC: American Psychological Association, 103–126.

Au KY (1999) Intra-cultural variation: Evidence and implications for international business. *Journal of International Business Studies* 30(4): 799–812.

Axelrod R (1997) The dissemination of culture: A model with local convergence and global polarization. *Journal of Conflict Resolution* 41: 203–226.

Axelrod R and Cohen MD (2000) *Harnessing Complexity: Organizational Implications of a Scientific Frontier*. Free Press: New York

Bandura A (1977) *Social Learning Theory*. New York: General Learning Press.

Banerjee A (1992) A simple model of herd behavior. *Quarterly Journal of Economics* 107(3): 797–817.

Barth F (ed.) (1969) *Ethnic Groups and Boundaries: The Social Organization of Culture Difference*. Boston: Little, Brown.

Beauvois J-L and Joule R-V (1996) *A Radical Dissonance Theory*. London: Taylor and Francis.

Bednar J (2009) *The Robust Federation: Principles of Design*. Cambridge: Cambridge University Press.

Bednar J and Page SE (2007) Can game(s) theory explain culture? The emergence of cultural behavior within multiple games. *Rationality and Society* 19(1): 65–97.

Bernheim D (1994) A theory of conformity. *Journal of Political Economy* 102(5): 841–877.

Bikhchandani S, Hirshleifer D, and Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy* 100(5): 992–1026.

Brock WA and Durlauf SN (2006) Identification of binary choice models with social interactions. *Journal of Econometrics* 140(1): 52–75.

Bouchaud J-P, Georges A, Koplik J, Provata A, and Redner S (1990) Superdiffusion in random velocity fields. *Physical Review Letters* 64: 2503–2506.

Bowers WJ (1968) Normative constraints on deviant behavior in the college context. *Sociometry* 31(4): 370–385.

Boyd R and Richerson PJ (1985) *Culture and the Evolutionary Process*. Chicago: The University of Chicago Press.

Calvert R and Johnson J (1998) Interpretation and coordination in constitutional politics. W. Allen Wallis Institute of Political Economy, University of Rochester, Working Paper 15.

Campbell DT (1965) Variation and selective retention in socio-cultural evolution. In: Barringer HR, Blankstein GI, and Mack RW (eds) *Social Change in Developing Areas: A Reinterpretation of Evolutionary Theory*. Cambridge, MA: Schenkman, 19–49.

Cavalli-Sforza LL and Feldman MW (1981) *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton, NJ: Princeton University Press.

Centola D, Gonzalez-Avella JC, Eguiluz VM, and San Miguel M (2007) Homophily, cultural drift, and the co-evolution of cultural groups. *Journal of Conflict Resolution* 51(6): 905–929.

Darmofal D (2005) Elite cues and citizen disagreement with expert opinion. *Political Research Quarterly* 58(3): 381–395.

Diamond J (2005). *Collapse: How Societies Choose to Fail or Succeed*. New York: Viking.

Durham WH (1991) *Coevolution: Genes, Culture, and Human Diversity*. Stanford, CA: Stanford University Press.

Durrett R and Levin S (2005) Can stable social groups be maintained by homophilous imitation alone? *Journal of Economic Behavior and Organization* 57(3): 267–286.

Festinger L (1957) *A Theory of Cognitive Dissonance*. Evanston, IL: Row and Peterson.

Fisman R and Miguel E (2007) Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy* 115(6): 1020–1048.

Friedkin NE (1984) Structural cohesion and equivalence explanations of social homogeneity. *Sociological Methods and Research* 12(3): 235–261.

Friedkin NE and Johnsen EC (1997) Social positions in influence networks. *Social Networks* 19(3): 209–222.

Gazzaniga MS (ed.) (1999) *Conversations in the Cognitive Neurosciences*. Cambridge, MA: The MIT Press.

Ginges J, Atran S, Medin D, and Shikaki K (2007) Moral barriers on rational resolution of violent political conflict. *Proceedings of the National Academy of Science* 104(18): 7357–7360.

Hannan TH (1979) Expense-preference behavior in banking: A reexamination. *Journal of Political Economy* 87(4): 891–895.

Harmon-Jones E and Harmon-Jones C (2002) Testing the action-based model of cognitive dissonance: The effect of action orientation on postdecisional attitudes. *Personality and Social Psychology Bulletin* 28(6): 711–723.

Henrich J (2000) Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *The American Economic Review* 90(4): 973–979.

Henrich J, Boyd R, Bowles S, et al. (2001) In search of homo economicus: Behavioral experiments in 15 small-scale societies. *The American Economic Review* 19(2), Papers and Proceedings of the Hundred Thirteenth Annual Meeting of the American Economic Association: 73–78.

Hofstede G (2001) *Culture's Consequences: Comparing Values, Behaviors, Institutions, and Organizations Across Nations*. Thousand Oaks, CA: Sage Publications.

Homans GC (1950) *The Human Group*. New York: Harcourt, Brace, and Company.

Huesmann LR (1988) An information processing model for the development of aggression. *European Journal of Personality* 3(2): 95–106.

Huesmann LR (1998) The role of social information processing and cognitive schema in the acquisition and maintenance of habitual aggressive behavior. In: Geen RG and Donnerstein ED (eds) *Human Aggression: Theories, Research, and Implications for Policy*. New York: Academic Press, 73–109.

Inglehart R (1997) *Modernization and Postmodernization: Cultural, Economic, and Political Change in 43 Societies*. Princeton, NJ: Princeton University Press.

Inglehart R and Welzel C (2005) *Modernization, Cultural Change, and Democracy: The Human Development Sequence*. Cambridge: Cambridge University Press.

Kameda T and Nakanishi D (2002) *Does Social/Cultural Learning Increase Human Adaptibility? Roger's Questions Revisited*. Japan: Hokkeido University.

Kennedy RJ (1988) The intermediary and social distance in Western Torres Strait. In: Mayers MK and Rath DD (eds) *Nucleation in Papua New Guinea Cultures. Dallas: International Museum of Cultures Publication* 23, 87–103.

Kitts JA, Macy MW, and Flache A (1999) Structural learning: Attraction and conformity in task-oriented groups. *Computational and Mathematical Organization Theory* 5(2): 129–145.

Klemm K, Eguíluz VM, Toral R, and San Miguel M (2003) Nonequilibrium transitions in complex networks: A model of social interaction. *Physical Review E* 67(2): 1–6.

Kuran T (1995) *Private Truths and Public Lies: The Social Consequences of Preference Falsification*. Cambridge, MA: Harvard University Press.

Kuran T and Sandholm W (2008) Cultural integration and its discontents. *Review of Economic Studies* 75: 201–228.

Liska AE (1975) *The Consistency Controversy*. New York: Wiley and Sons.

Macy MW, Kitts JA, Flache A, and Benard S (2003) Polarization in dynamic networks: A Hopfield model of emergent structure. In: Breiger R, Carley K, and Pattison P (eds) *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*. Washington, DC: The National Academies Press, 162–173.

March JG (1991) Exploration and exploitation in organizational learning. *Organization Science* 2(1): 71–87.

Marsden PV and Friedkin NE (1993) Network studies of social influence. *Sociological Methods Research* 22: 127–151.

McGuire WJ (1966) Attitudes and opinions. *Annual Review of Psychology* 17: 475–514.

McPherson M, Smith-Lovin L, and Cook JM (2001) Birds of a feather: Homophily in social networks. *Annual Review of Sociology* 27: 415–444.

Miguel E, Saiegh SM, and Satyanath S (2008) National cultures and soccer violence. NBER Working Paper number 13968.

Minturn L and Lambert WW (1964) *Mothers of Six Cultures: Antecedents of Child Rearing*. New York: Wiley.

Nisbett RE (2003) *The Geography of Thought: How Asians and Westerners Think Differently ... and Why*. New York: Free Press.

Nisbett RE and Ross LD (1980) *Human Inference: Strategies and Short-comings of Social Judgment*. Englewood Cliffs, NJ: Prentice-Hall.

North DC (1990) *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.

Page SE (1997) On incentives and updating in agent based models. *Economics* 10(1) 67–87.

Page SE (2007) *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton, NJ: Princeton University Press.

Page SE (2008) Uncertainty, difficulty, and complexity. *Journal of Theoretical Politics* 20(2): 115–149.

Page SE, Sander LM, and Schneider-Mizell CM (2007) Conformity and dissonance in generalized voter models. *Journal of Statistical Physics* 128(6): 1279–1287.

Pavlov IP (1903) Experimental psychology and psychopathology in animals. *Herald of the Military Medical Academy* 7: 109–121.

Pelto GH and Pelto PJ (1975) Intra-cultural diversity: Some theoretical issues. *American Ethnologist* 2(1), Intra-Cultural Variation: 1–18.

Roberts JM (1964) The self-management of cultures. In: Goodenough WH (ed.) *Explorations in Cultural Anthropology: Essays in Honor of George Peter Murdock*. New York: McGraw-Hill, 433–454.

Rogers EM (1995) *Diffusion of Innovations, Fourth Edition*. New York: Free Press.

Ross L and Ward A (1995) Psychological barriers to dispute resolution. In: Zanna M (ed.) *Advances in Experimental Social Psychology* 27. San Diego: Academic Press, 255–304.

Schelling TC (1971) Dynamic models of segregation. *Journal of Mathematical Sociology* 1: 143–186.

Sewell WH Jr (1999) The concept(s) of culture. In: Bonnell VE and Hunt L (eds) *Beyond the Cultural Turn. New Directions in the Study of Society and Culture*. Berkeley: University of California Press, 35–61.

Simmel G (1955) *Conflict and the Web of Group Affiliations*. New York: The Free Press.

Simon HA (1982) *Models of Bounded Rationality*. Boston: MIT Press.

Singer JE (1966) Motivation for consistency. In: Feldman S (ed.) *Cognitive Consistency*. New York: Academic Press, 47–73.

Skinner BF (1974) *About Behaviorism*. New York: Random House.

Thompson RW (1975) Gratification patterns in Buganda: An explanation of intra-cultural diversity. *American Ethnologist* 2(1): 193–206.

Tindall BA (1976) Theory in the study of cultural transmission. *Annual Review of Anthropology* 5: 195–208.

Tittle CR and Hill RJ (1967) Attitude measurement and prediction of behavior: An evaluation of conditions and measurement techniques. *Sociometry* 30(2): 199–213.

Van Overwalle F and Jordens K (2002) An adaptive connectionist model of cognitive dissonance. *Personality and Social Psychology Review* 6(3): 204–231.

Warner LG and DeFleur ML (1969) Attitude as an interactional concept: Social constraint and social distance as intervening variables between attitudes and action. *American Sociological Review* 34(2): 153–169.

Wedeen L (2002) Conceptualizing culture: Possibilities for political science. *American Political Science Review* 96(4): 713–728.

Weick KE (1969) *The Social Psychology of Organizing*. Reading, MA: Addison-Wesley.

Whiting BB (ed.) (1963) *Six Cultures: Studies of Child Rearing*. New York: John Wiley.

Young P (1998) *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press.

Young P and Burke M (2001) Competition and custom in economic contracts: A case study of Illinois agriculture. *American Economic Review* 91(3): 559–573.