APPENDICES

Cultural Diversity in Unequal Societies Sustained Through Cross-Cultural Competence and Identity Valuation

John A. Bunce

Department of Human Behavior, Ecology, and Culture
Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany.
Department of Anthropology, University of California, Davis, USA.
Email: john_bunce@eva.mpg.de

Appendix A. Other Dynamical Models of Cultural Diversity

Here I discuss a sample of influential models examining the effect of inter-group interaction on the dynamics of group-typical cultural norms, and I compare these with the model developed in the main text. A few of these models allow an individual to behave in coordination games in a way that differs from her own preferred behavior (what, in the main text, I call her preferred norm). The possibility of discordance between one's behavior and preferred norm is a prerequisite for cross-cultural competence (CCC). However, most of the models described below do not allow an individual to modify her behavior according to the behavior of her partner in a (correlative) coordination game. Rather, in most models, all individuals behave in one way with all of their interaction partners (though see section A.4, and Kuran and Sandholm 2008 for a case of limited flexibility). Thus, CCC in the context of coordination is not implemented (though see Kandler et al. 2010 for CCC without explicit coordination). A theme common to most of these models is that if/when behavioral (norm) diversity is maintained in a mixed equilibrium, individuals with different behaviors (norms) cannot coordinate. Similarly, in these models, any equilibrium in which all individuals can coordinate necessarily entails a loss of cultural diversity in the domain of coordination (though see exceptions in section A.4). Here, the context of interest is a hypothetical integrative society, which requires all individuals to be able to engage in mutually- (and optimally-, given exogenous constraints) beneficial interactions (coordination) in a particular behavioral domain. The question is whether a diversity of competing behaviors/norms in that domain can be maintained at equilibrium given such a context. All of the models described below yield important insights into cultural dynamics at group boundaries. However, as I show below, none of them can adequately address this specific question.

A.1. **Interaction Models with Mixed Equilibria.**

A.1.1. *Kuran and Sandholm (2008).* Kuran and Sandholm (2008) examined the dynamics of coordination behaviors in situations of initial behavioral diversity. In their model, each individual has an actual behavior x and a preferred behavior π, the values of each of which are drawn from a continuous range. Individuals are members of one of two groups, which have initially different distributions of behaviors and preferences. With probability c, individuals engage in a coordination interaction with a randomly-chosen individual in a border region containing both in- and out-group members. The proportion of in- and out-group members in this border region is an increasing function of each group's group-specific value of c and of its size m. With probability $1 - c$, individuals engage in a coordination interaction with a randomly-chosen in-group member in the group's core home region. It is assumed to be prohibitively costly for an individual to adjust her behavior according to the behavior of each of her interaction partners (their pg 203). Thus, individuals must choose one behavior to use with all interaction partners (both in- and out-group individuals) in the border region, and one, potentially different, behavior to use with all partners in the home region. Individuals are assumed to know the average behavior in both the in- and out-group. An individual's average utility for a coordination interaction in a given region is a decreasing function of the difference between her current behavior x and the average behavior of individuals in that region (this difference can be thought of as her coordination payoff/miscoordination cost), and a decreasing function of the difference between her current preferred behavior π and the average behavior of individuals in that region (this difference can be thought of as her personal payoff/cost). Individuals may weight these two payoffs unequally, such that ω is the group-specific weight given to coordination payoffs obtained through in-group interactions relative to personal payoffs, and a is the corresponding relative weight for out-group interactions. For a given interaction region, individuals choose a behavior that maximizes their average utility (their Equation 2). Thus, the behavior chosen by an individual is a function of her preferred behavior. Within a group, the distribution of preferences evolves to match the distribution of utility-maximizing (equilibrium) behavior, the rate of which is an increasing function of the difference between these two distributions (their Equation P). With respect to the evolution of behavior and preferences across both border and home regions, utility-maximizing equilibrium behavior in the border region is discounted by a factor $\lambda \in (0, 1]$ relative to equilibrium behavior in the home region (their pg 213).

Analysis of the Kuran and Sandholm (2008) model shows that behaviors and preferences in both groups evolve toward a single value π^*, which is an average of the initial average preferences in each group, weighted by relative group size and group-specific valuation of out-group coordination payoffs (a). Thus, cultural diversity (group-level variance in behavior and preferences) is inevitably lost from the population. A larger group, and a group with whom coordination is more highly valued by the other group, exerts more of an influence on the evolutionary trajectory, such that π^* is closer to this group's initial mean preference. In this model, cultural diversity can be maintained only by a constant influx of immigrants to each group, such that the initial preference distributions of the immigrants are independent of the distributions in the respective receiving groups (their section 5.2).

In the Kuran and Sandholm (2008) model, a, valuation of out-group coordination payoffs, is analogous to b in the model in the main text, the extra payoff benefit resulting from a successful coordination with a member of the out-group. In the model in the main text, the group with the lower relative b is considered to have higher power in the coordination interaction, as this group values out-group coordination less and thus wields a credible threat of voluntarily forgoing the interaction. Members of a group that receives a higher b are willing to sacrifice more (i.e., an amount up to the group-specific value b) in order to coordinate with out-group members. They thus have lower power in such interactions (Bunce and McElreath 2018). Provided low-power individuals sacrifice an amount less than b when coordinating with out-group members, they receive higher payoffs from out-group than from in-group coordination. Consequently, if the likelihood of out-group coordination is sufficiently high, norms which favor out-group coordination (i.e., out-group-typical norms) will increase in frequency among members of the low-power group. Thus, analogous to the model of Kuran and Sandholm (2008), groups that place greater value on coordination with the out-group tend to adopt out-group-typical norms to a greater extent than vice versa (Figure A.11B). Similarly, although relative group sizes are incorporated differently into the two models (group-specific parameters m in Kuran and Sandholm 2008, and a ratio of out-group interaction probabilities $\frac{1-a_L}{1-a_S}$ in the model in the main text), the overall effect is the same, namely, smaller groups are more likely to adopt out-group-typical norms than are larger groups (Figure A.11A).

Note that Kuran and Sandholm (2008) relate their parameter a indirectly to group-identity (their section 6.3). A minority group that values its cultural group identity will attempt to resist cultural convergence with the majority out-group. This resistance may manifest as a devaluation of payoffs obtained through coordination with the out-group, i.e., a low or negative value for a. This, in turn, moves the long-run behavior/preference equilibrium value π^* closer to the initial mean preference of the minority group, i.e., loss of minority-typical preferences is reduced. In contrast, in the model in the main text, group identity valuation i is explicitly defined as independent of the value placed on out-group coordination, $1 + b$. This facilitates an analysis of how strong group identity valuation must be in order to maintain minority-typical norms/preferences, given that the minority group values out-group coordination payoffs.

As a further contrast to the model of Kuran and Sandholm (2008), the model in the main text shows that, when cross-cultural competence is possible and valuation of group identity (i) is sufficiently high, group-level diversity of cultural norms can be sustained indefinitely, even in the absence of immigration. As explained in the main text, this is because cross-culturally competent individuals are able to receive the added benefits of out-group coordination while maintaining preferred norms typical of their in-group, given i sufficiently large to compensate for the cost of cognitive dissonance c.

A.1.2. *Carvalho (2017).* The model of Carvalho (2017) addresses the evolution of inter-group coordination behavior when individuals may engage in behavior that does not necessarily correspond with how they would prefer to behave. Like the model of Kuran and Sandholm (2008), this model permits discordance of behavior and preferences (norms), yet, at any given time, individuals are not permitted to tailor their behavior to that of their interaction partner. Thus, cross-cultural competence, as defined in the main text, is not implemented in this model. Furthermore, in contrast to the model in the main text, Carvalho (2017)'s model focuses exclusively on inter-group interaction. The model does not incorporate interaction among individuals within groups. Utility for individuals engaging in inter-group interactions is an increasing function of both coordination payoffs (if/when coordination is achieved), and personal payoffs for behaving in a certain way, regardless of whether or not such behavior results in coordination (his Equation 1). Thus, to achieve coordination, individuals may need to employ behavior that they do not prefer, i.e., behavior that does not yield them the highest possible personal payoff. Carvalho (2017) focuses on situations in which inter-group coordination is possible, and the payoffs to coordination always outweigh personal payoffs for employing a behavior that results in miscoordination (his Condition 2). In each time step, individuals may change their behavior to maximize their utility, given the current distribution of behaviors in the out-group (about which they have complete knowledge). There is a low probability ϵ that an individual chooses a behavior at random, regardless of its affect on utility.

Results of the model show that, when errors in behavior updating are present ($\epsilon > 0$), inter-group miscoordination can be a stable strategy, even if coordination yields higher payoffs than the maximum personal payoffs obtained through miscoordination. This is important for the present discussion because eliminating inter-group coordination eliminates the mechanism (in all models discussed here) by which cultural diversity is reduced. Thus, Carvalho (2017)'s model suggests that behavioral updating errors can preserve cultural diversity by reducing mutually-beneficial inter-group interaction. Furthermore, group-level differences in inter-group coordination payoffs can further destabilize inter-group coordination (his Equation 8). Such group-level inequality is analogous to group-level power differences in the model in the main text (where Carvalho 2017's Δ is analogous to b in the main text). In the Carvalho (2017) model, the group receiving the lower coordination payoff has less incentive to coordinate with the out-group, which decreases overall inter-group coordination. In the model in the main text, the group receiving the lower coordination payoff has higher power. As explained in Appendix section A.1.1, because members of this group have less incentive to coordinate with the out-group, the out-group is more likely to adopt this group's cultural norms.

In summary, cultural diversity can be preserved by mechanisms that decrease the incentives to individuals for inter-group interaction. Carvalho (2017)'s model shows that such mechanisms include behavior/norm updating errors, and structural group-level inequality. However, the model in the main text focuses on mechanisms to maintain cultural diversity in societies where mutually beneficial inter-group interaction, i.e., coordination, is desired (an "integrative" society).

A.1.3. *Advani and Reich (2015).* The model of Advani and Reich (2015) investigates the dynamics of social configuration in a society comprising a majority and a minority group. Initially, all majority members have "cultural" behavior (or, equivalently, norm) x_M and "non-cultural" behavior y_M, and all minority members have behaviors x_m and y_m. Each individual decides whether or not to form an interaction tie with each other individual in the population. Each tie formation comes with a cost L, suffered by the initiator of the tie. An individual's utility is the sum, over all of her ties, of the payoffs she receives from coordination with each alter on the cultural behavior and on the non-cultural behavior (with weight $\alpha \in [0, 1]$ given to the cultural coordination payoff), minus the cost L of forming each tie. From this sum, a cost c is subtracted if the individual changes her cultural (but not non-cultural) behavior prior to forming all interaction ties (their Equation 2). Individuals choose their behavior in order to maximize utility, and they have complete information regarding the behaviors of all individuals in the population. Each individual must choose one behavior to employ in interactions with all of her alters, and thus this model does not incorporate CCC as conceived in the main text.

Analysis of the model reveals that, if the minority group is sufficiently large relative to the majority group, two forms of mixed equilibia are possible in which both minority- and majority-typical cultural norms are sustained. In a "multicultural" mixed equilibrium, members of each group (minority and majority) retain their respective cultural behaviors. However, they adopt a common non-cultural behavior, and each person establishes interaction ties with everyone else in the population. In a "segregation" mixed equilibrium, members of each group retain both their initial cultural and non-cultural behaviors, and establish interaction ties only with members of their respective in-group. Both types of mixed equilibria can be achieved with smaller minority groups as the cost of changing cultural behaviors, c, increases (such that the minority is less willing to adopt majority cultural behavior). As the value placed on cultural (relative to non-cultural) coordination payoffs, α increases, and as the cost of forming a tie, L, increases, larger minority group sizes are necessary to facilitate multi-cultural equilibria, and smaller sizes are needed for segregation equilibria. In other words, a multi-cultural equilibrium (many cross-group ties and shared non-cultural behavior) is more likely when more value is placed on non-cultural coordination benefits, and when forming ties is easier. In contrast, if the minority group is sufficiently small relative to the majority group, minority-typical behavior will be lost from the population. In this case, all minority individuals adopt both the cultural and non-cultural behaviors of the majority group, and form ties with everyone else in the population.

In Advani and Reich (2015)'s model, the value attributed to payoffs resulting from coordination on "cultural" behaviors (α) shares similarities with i in the model in the main text (group identity valuation), in the sense that both of these parameters contribute (indirectly) to the valuation of group-typical behaviors/norms. An important aspect of the dynamics of Advani and Reich (2015)'s model, namely the formation of multicultural equilibria, is attributable to the assumption that the valuation of "cultural" coordination payoffs trades off against the value of payoffs from coordination on "non-cultural" behaviors $(1-\alpha)$. Thus, when a minority group is relatively large, increasing α tends to decrease the likelihood that inter-group interaction ties will form (e.g., none will form if $1 - \alpha < L$, their Proposition 1). In the model in the main text, inter-group interaction is exogenously imposed. Incorporating Advani and Reich (2015)'s tie formation mechanism may be a fruitful extension.

In one interpretation of Advani and Reich (2015)'s model, a multicultural mixed equilibrium could entail the loss of minority-typical behaviors/norms in some domains ("non-cultural" behaviors). This loss then facilitates inter-group coordination in those domains. However, minority-typical behaviors/norms are preserved in other domains ("cultural" behaviors), rendering inter-group coordination in such domains impossible. A multi-cultural mixed equilibrium where all individuals interact with all others occurs when the coordination payoffs in the domain of minority-typical behavioral loss are sufficiently high to compensate for the cost of initiating an interaction, given the absence of coordination payoffs in domains where group-level behavioral diversity is preserved. Thus, like most other models that exclude CCC, the benefits of inter-group interaction trade off against the benefits of maintaining group-typical norms (or behaviors). This tradeoff certainly seems plausible for many aspects of inter-group interaction. As Advani and Reich (2015) state, religious strictures may prevent adherents of two different faiths from coordinating in certain domains ("cultural" behaviors, in their model), as coordination would require members of one group to violate their strictures. Diversity in such religious domains is maintained at the cost of coordination. However, religious strictures may not apply in other domains, such as sports, and individuals of one faith may adopt the recreational behavior of the other, thereby facilitating coordination and the concomitant loss of behavioral diversity in this domain ("non-cultural" behaviors, in their model). If the benefits of coordinating in the domain of sports are higher than the cost of initiating an inter-group interaction, a possible outcome is that everyone interacts with everyone else (in the domain of sport), yet religious diversity is maintained.

In contrast, the model in the main text is applicable to contexts in which coordination and the maintenance of cultural diversity are both desired in the same domain. For instance, two groups within a society may differ in their norms of behavior in educational settings, with regard to gender roles, in their dealings with authority figures, or in the context of a fair exchange of resources. The model in the main text, incorporating CCC and the valuation of group identity, illustrates one way in which members of the two groups could engage in mutually beneficial interactions with each other in these particular domains while maintaining their respective norms.

A.1.4. *Bunce and McElreath (2018).* Bunce and McElreath (2018) modelled norm dynamics in a society comprising a disempowered minority group and a powerful majority group, each with initially different distributions of coordination norms. Each individual employed only a single norm during all interactions in a given time step. Therefore CCC was not implemented. This model is designed to address contexts in which minority and majority group members each have separate non-overlapping territories, and a fraction m of each group visits the other group and engages in interaction. Individuals have complete information about the norms held by all others, and they have an ability a to choose an interaction partner who has a norm matching theirs (as opposed to choosing a partner at random with respect to norm). The relative power of groups is defined exactly as in the model in the main text.

Analysis of Bunce and McElreath (2018)'s model under plausible parameter conditions shows that mixed equilibria are possible when individuals cannot assort accurately on norm (low a), and members of the powerful majority group do not visit the territory of the disempowered minority, but minority group members do visit the territory of the majority (their Figures 2C, S3C, and S4F). However, at such mixed equilibria, interaction between people holding different norms yields sub-optimal payoffs for both of them relative to interaction with people holding the same norm. Thus, one might expect people in such a society to quickly segregate themselves by norm (perhaps through the evolution of markers: (McElreath et al. 2003)). In contrast, in an integrative society, in which all members are expected to engage in mutually beneficial interactions with all others, such a mixed equilibrium would be unsatisfactory. In the model in the main text, the implementation of CCC facilitates coordination among all individuals at mixed equilibria.

A.1.5. *Kandler et al. (2010).* Kandler et al. (2010) modelled the process of language shift in a society comprising speakers of a high-status and low-status language. Note that language is an example of a set of cultural norms, and inter-individual communication is a coordination interaction. Their model incorporated the mobility of speakers, demographic dynamics of the language groups, and bilingualism (one dimension of CCC, as defined in the main text). Whenever two monolingual speakers of different languages come into contact, they each have a probability of becoming bilingual. This probability is higher for the speaker of the lower status language. Similarly, whenever a bilingual speaker comes into contact with a monolingual speaker, the bilingual speaker has a probability of becoming monolingual. This probability is lower when the monolingual individual speaks the low status language (their Figure S2). Analysis of this model reveals that no mixed equilibria are possible, and thus one language (usually, but not always that with low-status) is inevitably lost from the population.

However, Kandler et al. (2010) modify their original model by implementing a context in which only the high status language must be spoken, and another context in which only the low-status language must be spoken, conditional on there being a sufficient number of bilingual speakers (their Diglossia model). With this modification, mixed equilibria that include bilingualism become possible. As Kandler et al. (2010) point out, exogenously-imposed language revitalization programs are one way to create a need for a low status language, and thereby ensure its continued maintenance in the form of bilingual individuals.

This model differs in several important ways from the model in the main text. For instance, in Kandler et al. (2010)'s model, an individual's group identity is synonymous with her language (equivalently, norm). Thus, when language changes, group identity changes. Similarly, individuals lose their original group identity when they become bilingual (equivalently, CCC). For this reason, the parameter i (valuation of group identity), and the four different CCC phenotypes ($S1X$, $S2X$, $L1X$, and $L2X$) in the model in the main text have no analogues in the Kandler et al. (2010) model. Consequently, a different mechanism is required in their model to maintain language (cultural) diversity in the population at equilibrium, namely, (the external imposition of) a context in which only the low status language can be spoken.

A.2. **Immigration Models with Mixed Equilibria.**

A.2.1. *Boyd and Richerson (2009).* Boyd and Richerson (2009) develop a model to investigate the effect of payoff-biased migration on the evolution of group-typical behaviors (equivalent to norms, as modeled in the main text) that differ in the coordination payoffs they provide. Individuals may hold only a single norm at a time, and they use this norm in interactions with all others. Thus, CCC is not implemented. In this model, if the rate of migration is low relative to the strength of payoff-biased copying within groups, then groups can evolve mixed equilibria, such that both the higher coordination payoff-yielding behavior/norm and the lower coexist (their Figures 5-7). Thus, if a subgroup is considered to be a society, a constant small number of immigrants with a norm different from that of the receiving society, coupled with payoff-biased updating of norms, can contribute to the maintenance of cultural diversity within the society. However, analogous to the model of Bunce and McElreath (2018), at such mixed equilibria immigrants and residents with different behaviors/norms cannot engage in mutually-beneficial interaction, and would be expected to segregate. This would likely be unsatsifactory from the perspective of an integrative society. In contrast, if immigrants became cross-culturally competent, perhaps through a mechanism like that proposed in the model in main text, then interactions with residents would entail coordination.

A.2.2. *Mesoudi (2018).* Mesoudi (2018) develops a model to investigate the effect of migration and conformist social learning on the maintenance of between-group cultural diversity. Here, again, CCC is not implemented, as individuals hold only one norm at a given time. Analysis of the model shows that, in any one sub-group, the rate of immigration of individuals with different group-typical norms and the strength of conformist social learning can balance to facilitate mixed equilibria, in which both norms are maintained in the sub-group (e.g., his Figures 1 and 4). Thus, complementing the model of Boyd and Richerson (2009), a constant small number of immigrants with a norm different from that of the receiving group, coupled with conformist social learning of norms (from, ideally, a large number and randomized composition of demonstrators), can contribute to the maintenance of cultural diversity within a group. Also analogous to Boyd and Richerson (2009), in the absence of CCC, such mixed equilibria in an integrative society would be expected to be unsatisfactory, likely resulting in either segregation or high levels of mutually sub-optimal inter-personal interactions.

A.2.3. *Erten et al. (2018).* The model of Erten et al. (2018) investigates the effect of migration and cultural conservatism on the sustainability of cultural diversity within groups. The focal group contains residents, all of whom hold the same norm r. At each time step, there is a constant influx of immigrants who all hold norm i ($\neq$ norm r). CCC is not implemented, as residents and immigrants hold only one norm at a time. Individuals have a probability X_{ir} of choosing to interact with individuals holding a norm different from their own. In the coordination version of their model, an individual's payoff in a given time step is directly proportional to the probability that she has the opportunity to, and chooses to, interact with another individual who holds her norm. An individual's decision to update her norm in a given time step is payoff-biased, but the probability of adopting a different norm is decreased by a factor of c, the individual's cultural conservatism (unwillingness to change norms).

Results suggest that low levels of migration, a high propensity to interact with others who hold different norms (high X_{ir}), and higher cultural conservatism (c) on the part of residents compared to immigrants, can result in a stable mixed equilibrium such that both resident-typical and immigrant-typical norms are sustained in the focal group (their Figure 4). The focal group in this model is analogous to the population in the model in the main text, consisting of both groups S and L. Residents in the Erten et al. (2018) model are comparable to minority-group members (group S) in the simplified model in the main text, who are exposed to out-group members (group L) whose initial norm distribution is unaffected by norm dynamics within the minority group. Similarly, cultural conservatism is analogous to group identity valuation (i) in the model in the main text, increasing the value of which can contribute to the sustainability of cultural diversity in both models.

Like the models of Boyd and Richerson (2009) and Mesoudi (2018), a mixed equilibrium in the coordination version of the Erten et al. (2018) model would seem to be an unsatisfactory out-come from the perspective of an integrative society which promotes mutually beneficial (coordination) interactions among all of its constituents. Although Erten et al. (2018) do not explicitly model payoffs from individual-level coordination interactions, one interpretation of a mixed equilibrium in this model (in the absence of CCC) is that resulting interactions between individuals holding norms i and r yield sub-optimal payoffs for both individuals. Importantly, to achieve a mixed equilibrium, individuals must be willing to interact with others who hold norms different from their own (high X_{ir}, their Figure 4). Thus, if given a choice, it seems that individuals in such a mixed equilibrium would segregate themselves by norm in order to ensure reliable coordination payoffs. In contrast, in the model in the main text, incorporation of CCC, as well as incentives to interaction with out-group members (b), facilitate both mixed equilibria at the society level (including both groups S and L) and mutually beneficial coordination among all individuals at such equilibria.

A.3. **Other Models.**

A.3.1. *Bisin et al. (2011).* Bisin et al. (2011) investigate the evolution of "oppositional" identities in a population consisting of a majority and minority group. An oppositional identity characterizes members of the minority group who reject the behavior/norm most common in the majority group because they wish to distinguish themselves from the majority. The maintenance of oppositional identity among (at least some) members of the minority group is equivalent, in the framework of the model in the main text, to the maintenance of cultural diversity in the population. In the Bisin et al. (2011) model, a minority individual chooses both a behavior (equivalently, a norm), either “mainstream” like the majority group or “oppositional”, and the degree to which she values the identity associated with that behavior (her α). Note that identity here is associated with an individual’s behavior, rather than with her minority/majority group affiliation (i.e., α in this model is not equivalent to i in the model in the main text). An individual’s choice of behavior is a function of the effort (τ) devoted by her parents to ensuring that she learn the behavior that they have. An individual chooses the degree to which she values her identity in order to maximize her utility. Utility is a function of coordination payoffs and the personal payoffs that one receives from performing a given behavior. An individual chooses a single behavior and a single degree of identity valuation for her entire lifespan. Therefore, CCC is not implemented in this model, and, like most other models described here, the maintenance of cultural diversity in such a context necessarily entails considerable miscoordination. Dynamics in the Bisin et al. (2011) model occur on the scale of generations, and are driven by the effort parents devote to ensuring their children adopt the behavior of the parents, i.e., parents’ socializing effort (τ). A key feature of the dynamics is that parents devote less effort to socializing their children as the probability that children would learn the parents’ behavior anyway by copying a randomly-chosen role model increases. This is termed “cultural substitutability”.

Analysis of this model suggests that oppositional identity among the minority group can be maintained at equilibrium if children tend to choose role models from their in-group, the minority group is large, the personal payoff advantage to performing the oppositional behavior is large, and/or interaction between members of minority and majority groups is frequent. Note that this last prediction seems to conflict with a prediction of the model in the main text, namely that, given a fixed level of group identity valuation (i), the sustainability of a minority-typical norm is less likely as individuals engage in more inter-group interaction (i.e., high a, Figure 2b).

To understand these conflicting results, note that, in the Bisin et al. (2011) model, individuals increase their valuation of (behavioral, rather than group) identity (their α) as the risk of miscoordination increases (their α^*, Equation 6). For minority individuals with the oppositional behavior, this risk increases as the probability of interacting with members of the majority group (all of whom have the mainstream behavior) increases. In this model, the larger a parent's α, the more likely she is to teach her child her behavior. This mechanism can sustain the minority oppositional behavior across generations. In contrast, in the model in the main text, individuals cannot strategically choose their group identity valuation (i) in order to compensate for the risk of miscoordination or cognitive dissonance (c) associated with the minority-typical norm. Thus, as the risk of miscoordination or cognitive dissonance increases (e.g., due to more inter-group interaction), individuals' average payoffs decrease, and norm (equivalently, behavior) updating is biased toward norms yielding higher average payoffs. If and how individuals can strategically modify their valuation of identity is a open question that requires further study.

A.3.2. *Olcina et al. (2018).* Olcina et al. (2018) investigate the effects of network structure on the dynamics of norms. They model a minority group in which each individual has a personal norm s within the continuous range $[0, 1]$, and performs an action $x \in [0, 1]$ that may or may not match her personal norm. In their description of the model, s is framed as a person's preferred degree of adoption of a majority norm, and action x is her actual degree of adoption of the majority norm. Individuals interact with a subset of others within the minority group to whom they have established (directed) ties. At each time step, an individual suffers a utility cost that increases with the distance between her chosen action x and the average action of the other group members with whom she has ties (social interaction payoff), and another utility cost that increases with the distance between her chosen action x and her preferred action s (personal payoff). The parameter ω represents the degree to which she values the social versus the personal payoff (their Equation 1). Individuals choose their action x at each time step in order to minimize their overall utility cost. After action x is chosen, an individual then updates her preferred action s. The parameter $\gamma \in [0, 1]$ represents the degree to which her preferred action is updated to match her actual action, as opposed to remaining unchanged from the previous time step. Note that, in this model, although an individual's performed action may not correspond with her preferred action (her norm), she must perform the same action with all of her interaction partners in any given time step. Thus, CCC is not implemented.

Analysis of the model suggests that, at equilibrium, the action x and the norm s of each individual converge to the same value. However this value may differ between individuals who are in different parts of the overall minority group social network. The equilibrium value $x = s$ of an individual is a function of the average initial norm value of the other individuals with whom she has interaction ties, weighted by a measure of the connectedness (eigenvector centrality) of each of these interaction partners. Thus, within the minority group, individuals belonging to different interaction subgroups (communication classes) that are not connected to each other may converge on different norms, resulting in the maintenance of norm diversity within the population. However, if the social network is strongly connected, e.g., each minority individual is either directly or indirectly connected to every other minority individual, then the norms of all individuals converge to the same value at equilibrium, and diversity is lost.

Olcina et al. (2018)'s model demonstrates the importance of network effects for the dynamics of norms. Individuals, or subgroups, that are isolated from interaction with the rest of the population can maintain distinctive norms at equilibrium. This is analogous to the finding, in the model in the main text, that distinctive minority norms are more likely to be maintained at equilibrium if individuals have a high affinity (a) for interactions with in-group members (Figure 2b). In the model in the main text, we are particularly interested in situations in which there is a high level of both in-group and out-group interaction (e.g., $a \approx 0.5$). Therefore, the mechanism that preserves norm diversity in the Olcina et al. (2018) model is less likely to play a major role. However, accounting for the network structure of inter-group interactions is undoubtedly important in the analysis of empirical data relating to norm dynamics, as suggested in Appendix B.1.1.

A.4. **Competing Norms vs. Complementary Coordination.** In the models described above (including the model in the main text), interactors receive a high payoff if they hold the same norm, and a low payoff if they hold different norms. This is termed correlative coordination (O'Connor 2019), and the two alternative norms can be thought of as mutually-exclusive and in competition with one another. As shown above, it appears that, without cross-cultural competence, any mixed equilibria that evolve under a regime of correlative coordination will necessarily entail either segregation of interactors or some degree of miscoordination in the population. Both of these outcomes is suboptimal from the the perspective of an integrative society.

However, researchers investigating the evolution of inequality commonly employ a class of model that allows mutually beneficial intergroup interaction under conditions of norm diversity, with or without cross-cultural competence. These models usually have a payoff structure that O'Conner (2019) terms complementary coordination. In contrast to correlative coordination, complementary coordination occurs when interactors receive a high payoff if they hold different norms and a low payoff if they hold the same norm. The alternative norms complement each other. Complementary coordination is a useful representation of situations like division of labor, where interactors both receive a greater benefit if they perform complementary behaviors rather than the same behavior. Intuitively, with such a payoff structure, it is usually much easier maintain a diversity of behaviors/norms at equilibrium than it is under correlative coordination. For researchers investigating the evolution of inequality, a particularly interesting situation arises when two interactors each receive a different payoff from complementary coordination. Under many such conditions, stable mixed equilibria can evolve such that different behaviors/norms become common in different groups, and inter-group interaction is therefore mutually beneficial, but results in payoffs distributed unequally by group (O'Connor 2019).

I would argue that complementary coordination does not seem like a natural way to represent inter-group interaction at most cultural boundaries where the question of interest is the sustainability or loss of group-typical norms. In such cases, the norms whose dynamics are of interest tend to be mutually-exclusive (competitive) rather than complementary, and interaction is thus better represented as correlative coordination. Complementary coordination is certainly extremely common at boundaries between cultural groups, but I propose that it is possible only after an equilibrium has been reached via evolution under a correlative coordination payoff regime. My reasoning is that complementary coordination requires interactors to assume complementary roles. However, this is only possible if both interactors agree on what those roles entail, and, indeed, if they agree that roles exist at all. Belief in the existence of a given set of roles is itself a norm that is mutually exclusive with respect to a norm for the existence of different (or no) roles. Interactors must first coordinate (correlatively) on a norm for a given set of roles, and only then can they coordinate (complementarily) on which of them plays which role (see also discussion in O'Connor (2019), pgs 21–22).

For example, in Matsigenka society, women grow, spin, dye, and weave cotton (ampei) into tunics (magatsi). There is a norm such that each adult Matsigenka woman should attain competence in each step of this process because, when she gets married, she will be in charge of making clothes for her husband and children. Up until about 60 years ago, Matsigenka families were relatively mobile and often highly dispersed, such that a woman could not always count on the presence of other adult women to help her make tunics. (Note that such norms related to Matsigenka familial economic independence are in the process of changing.) Call this norm for competence in every step of the tunic-making process the Generalist norm. Contrast this with a Specialist norm, which holds that each woman should become competent in only a subset of the skills needed to make a tunic. For instance, under the Specialist norm, some women would specialize in spinning and others in weaving. Now suppose there is a situation where there are two adult sisters whose elderly parents' house burned down, destroying the tunic of each parent. The two sisters must make a total of two new tunics for their parents, and they must decide how to divide up the labor. Each sister may hold either a Generalist norm or a Specialist norm. Conditional on her holding a Specialist norm, she may hold either a Spinning norm or a Weaving norm. The sisters must first play a correlative coordination game with the competing norms Generalist and Specialist. If they hold different norms in this game, then at most only one tunic will be made (by the Generalist sister) and they will not have solved the problem. If both sisters hold the Generalist norm, then each makes one tunic from beginning to end and the problem is immediately solved. However, if both sisters hold the Specialist norm, then they must play a complementary coordination game with the norms Spin and Weave. In this game, if both sisters hold the same norm, no tunics are made. If they hold different norms, then they can combine their skills to make two tunics and the problem is solved.

The models described in the previous sections, along with the model in the main text, use a correlative coordination payoff structure to represent dynamics among competing norms in situations analogous to, for example, inter-ethnic interaction between Matsigenka, most of whom may hold the Generalist norm, and Mestizos, most of whom may hold the Specialist norm. The question of interest is, can these two competing norms ever both be present at a stable equilibrium in which inter-group interaction is both common and mutually-beneficial? In contrast, models using a complementary coordination payoff structure would be a more appropriate representation of norm dynamics in a situation in which the Specialist norm has already gone to fixation in both the Matsigenka and Mestizo groups, and dynamics now involve only the complementary norms Spinning and Weaving. If inter-group interaction is frequent, each of these complementary norms may go to fixation in a different group, but neither is likely to go to fixation in the entire population. Dynamics of the Spin and Weave norms would be outside the scope of the model in the main text.

With this in mind, it is important to recognize that many models of complementary coordination include cross-cultural competence, in the sense that individuals are capable of learning each of a set of complimentary norms, and employing one norm during coordination with in-group members, and another when interacting with out-group members. In one such two-strategy model, Hoffman (2006), finds that, at all stable equilibria, both complementary norms occur at equal frequencies within groups (group = specific combination of marker traits). However, different groups often use different norms for out-group interactions. Rubin and O'Conner (2018) use a model with a three-strategy Nash demand game, which involves a combination of correlative and complementary coordination (interactors may coordinate correlatively on the Med norm, or complementarily on the High and Low norms). At most equilibria, a single norm (Med) went to fixation in all groups with respect to in-group interactions. However, depending on model structure and parameter values, equilibria were often mixed with respect to the norms used during out-group interactions. Note that both of the two previous models incorporate CCC. However, Henrich and Boyd (2008) and Erten et al. (2018, "complementation scenario", their Figure 4c,d) demonstrate that, in models with complementary coordination, stable mixed equilibria may be likely even in the absence of CCC. The Hoffmann (2006), Rubin and O'Conner (2018), and Henrich and Boyd (2008) models are designed to investigate the evolution of social stratification and inequality, rather than the sustainability and loss of group-typical norms. I have included such models of complementary coordination here for completeness, because they demonstrate equilibria where mutually-beneficial inter-group interaction can co-evolve with a diversity of norms, with or without CCC. However, as argued above, a correlative (rather than a complementary) coordination payoff structure is a more appropriate representation for the processes under investigation in the main text, namely the effect of inter-group interaction on the dynamics of competing group-typical cultural norms.

Appendix B. Detailed Methods

B.1. **Empirical Evidence of Cross-cultural Competence.** Measuring cross-cultural competence in a real-world population is a nontrivial task, as there is consensus on neither a definition (Spitzberg and Changnon 2009) nor a measurement strategy (Fantini 2009) for cross-cultural competence. In what follows, I present one definition and a new instrument, which I then use to measure the distributions of the various forms of uni- and cross-cultural competence in an Amazonian population consisting of interacting members of a minority indigenous Matsigenka group and a majority Mestizo group. I show that these distributions appear to vary with the degree and form of inter-group interaction in ways consistent with predictions of the theoretical model developed in the main text and below.

B.1.1. *Defining and Measuring CCC.* Following previous work (Bunce 2020), I define cross-cultural competence (CCC) as knowledge of, and the ability and willingness to use, norms typical of both the subjectively-defined in-group and out-group. Thus, a proxy (necessary but not sufficient condition) for cross-cultural competence is knowledge of the (potentially very different) distributions of norms in the in-group and out-group. The degree of such knowledge may vary by norm, such that an individual may be cross-culturally competent for one particular norm (e.g., she knows that the greeting norm of bow versus handshake differs between groups) but not another (e.g., she doesn't know that the eating norm of chopsticks versus fork differs). An individual's knowledge of the distributions of a particular norm can be assessed by asking her to guess the proportion of in- and out-group members who prefer each norm (e.g., bow or handshake). Guesses can then be compared to the actual distributions of the norms preferred in each group, and categorization of an individual as uni- or cross-culturally competent, according to this proxy, is a function of the accuracy of her guesses. If the distribution of a norm differs markedly between groups, such that most members of one group prefer it, while most members of the other prefer the alternative, then a norm can be conceived of as either "typical" of the in-group or typical of the out-group. When this is the case, and under this definition, two forms of cross-cultural competence are possible: 1) individuals prefer the norm typical of the in-group, yet know the norm typical of the out-group; and 2) individuals prefer the norm typical of the out-group, yet know the norm typical of the in-group. In the theoretical model, these two forms are represented by the phenotypes $S1X$ and $S2X$, respectively, for members of minority group S.

Previous work (Bunce 2020) used a relative measure to compare the degree of cross-cultural competence of different types of people within a society. However, in order to more directly address the predictions of this theoretical model, I here calculate an absolute measure of individual-level cross-cultural competence. Such a measure requires an additional set of assumptions, which I now explain.

Asking participants to guess the proportion of people in each group who prefer a given norm may be extremely challenging in societies in which people are unfamiliar with Western representations of proportions. However, it is usually easier to ask them to guess the norm most commonly preferred in each group. The absolute measure of individual-level cross-cultural competence developed here is designed for situations in which a different norm is preferred by a majority of each group. In this case, an individual's guesses are unambiguously either correct or incorrect, and her own preferred norm is either typical of her in-group or typical of the out-group (but not both). Eight combinations of preferred norms and guesses are possible. Using the definition above, four combinations of response patterns are easily classified: two uni-culturally competent ($S11$ and $S22$), and two cross-culturally competent ($S1X$ and $S2X$), as shown in Table A.1. However, as shown in the table, the other four combinations of individual responses are more difficult to classify. These individuals have knowledge of a norm different from that which they prefer (i.e., they are not uni-culturally competent). However, they are mistaken about which norm is most common in one of the two groups ($S1in$ and $S2out$), or in both groups ($S1both$ and $S2both$). Therefore, they are not cross-culturally competent either. As an overly-simple heuristic, phenotypes $S1in$ and $S2out$ in this context could be thought of as false pariahs, i.e., they mistakenly believe that most people in both the in- and out-group prefer a norm different from that which they prefer. Phenotypes $S1both$ and $S2both$ could be thought of as inaccurate observers, i.e., their belief about which preferred norm is most prevalent in each group is precisely the opposite of reality.

One plausible explanation for the existence of these unexpected phenotypes is the fact that individual social networks may be non-random subsets of both in- and out-group individuals. Thus, the perceived frequencies of norms preferred by people in any one individual's social network may be very different from the frequencies of preferred norms that she would perceive had she the opportunity to observe a random sample of the entire population of in- and out-group members. Cross- and uni-cultural competence may be best determined using the perceived frequencies of preferred norms within the subset of in- and out-group members with whom an individual is likely to interact, i.e., her social network. Thus, a phenotype that is cross-culturally competent within one person's social network may be different from the phenotypes that are cross-culturally competent for an individual whose social network potentially includes the entire population of in- and out-group members (i.e., the situation assumed in the theoretical model above).

For this reason, individuals with the phenotypes $S1in$, $S2out$, $S1both$, $S2both$ may well be cross-culturally competent within their personal social networks. Alternatively, empirical observation of these phenotypes may simply be an artifact of the methods used to determine phenotype (e.g., participants not understanding interview questions in the way intended). For instance, the particularly high frequencies of the phenotypes $S2out$ among Matsigenka and $L1out$ among Mestizos (Figures A.3 - A.6) may result from the fact that these interviewees generally believe both in- and out-group members prefer the same norms that they do. However, they changed their mind about the norm that they personally prefer during the time interval (one week to over a year: Bunce (2020)) between the personal norm interview and the guessing interview, or they interpreted the question in a different way the second time around.

Table A.1. Phenotype classification for group S, where the majority of the in-group prefers norm 1 and the majority of the out-group prefers norm 2

Own Preferred Norm	Guess for In-Group[a]	Guess for Out-Group	Phenotype Classification[b]
1	1	1 (W)	$S11$
2	2 (W)	2	$S22$
1	1	2	$S1X$
2	1	2	$S2X$
1	2 (W)	1 (W)	$S1both$
1	2 (W)	2	$S1in$
2	1	1 (W)	$S2out$
2	2 (W)	1 (W)	$S2both$

[a]Incorrect guesses are indicated by (W)

[b]Uni-culturally competent phenotypes are $S11$ and $S22$. Cross-culturally competent phenotypes are $S1X$ and $S2X$. Other phenotypes are explained in the text.

B.1.2. *Limitations of the Metric.* This metric of absolute individual-level cross-cultural competence should be regarded as a first approximation, as it has several limitations apart from the complexities of classification shown in Table A.1. As described above, the metric is designed for contexts where the majority of each group prefers a different norm. However, there may be situations of interest in which a norm is almost completely absent in one group, but preferred by a non-trivial minority of the second group (e.g., some of the norms studied by Bunce 2020). An individual who knew the true distributions of the norm in each group and an individual who believed the norm to be completely absent in both groups would both be classified as cross-culturally competent using this metric.

Additionally, the metric does not scale with the difficulty of making a correct guess. For instance, if the true proportion of people who prefer norm 1 is 0.49 in one group and 0.51 in the other group, it may be very difficult to guess which norm the majority of people in each group prefer. Thus, even if a participant is reasonably knowledgeable about the norm distributions in both groups, there is a high probability that she will guess incorrectly and thus not be classified as cross-culturally competent. In contrast, the same reasonably knowledgeable participant would be much more likely to make correct guesses if the norm distributions in the two groups were 0.1 and 0.9, respectively.

As noted in the above definition, cross-cultural competence is often conceptualized as requiring more than just the knowledge of norms measured by the current metric. It is also thought to entail an ability and willingness to interact using those norms. For instance, an individual who knows that chopsticks are the norm in one society, but insists on using a fork when visiting, might not be considered cross-culturally competent. However, the metric developed here, which relies only on knowledge, would classify her as such.

The present individual-level absolute metric is developed for the sole purpose of demonstrating the plausibility of the theoretical model, and thus it may be of limited use beyond this context. The experience-based relative measure of cross-cultural competence developed in previous work (Bunce 2020) is designed to be more widely applicable, though it is less amenable to direct comparisons with the simplistic theoretical model presented in the main text. It is hoped that the model and metric described here will inspire future efforts to better bridge the existing gap between theoretical and empirical study of the population-level consequences of cross-cultural competence.

With these limitations in mind, below I describe implementation of the metric in an Amazonian population consisting of minority indigenous Matsigenka and majority Mestizos.

B.1.3. *Empirical Methods Overview.* All data are taken directly from Bunce (2020), where the methodology is described in detail. An overview of data collection is provided here.

An interview comprising fourteen ethnographically-informed vignette questions measured preferred cultural norms across a range of interaction domains among 74 adult residents of the Matsigenka Native Community of Tayakome and 84 residents of the neighboring Mestizo towns of Boca Manu and Atalaya in the Manu region of the Department of Madre de Dios, in Amazonian Peru (Bunce and McElreath 2017; Shepard et al. 2010; Llosa Isenrich and Nieto Degregori 2003). Note that, as detailed in Bunce (2020) (see his Table 2), there is a high, but imperfect, correspondence between residence and self-identified ethnicity among participants. In-group and out-group categories for each participant are determined on the basis of self-identified ethnicity. A subset of 53 participants from the Matsigenka community and 50 participants from the Mestizo communities was then asked to guess the most common preferred norm (i.e., the response of a randomly-chosen member) in their own ethnic group (in-group) and in the other ethnic group (out-group) for each vignette question, similar to previous methods for measuring inter-group perceptions (Medin et al. 2007; Gurven et al. 2008). Based on ethnographic observations collected over a year (Matsigenka) and five months (Mestizos) of participation in community life, domains of salient inter-ethnic interaction (e.g., education, labor, commerce) were identified and interviewees' self-reported experience in each domain was recorded. Data were analyzed using Bayesian estimation of item-response theory (IRT) models (Bunce and McElreath 2017), which resulted in posterior distributions (i.e., model estimates with associated uncertainties) of the probabilities of preferring particular norms and guessing the most common norms preferred by in-group and out-group members, for each experience type. From this, Bunce (2020) calculated a relative measure of cross-cultural competence and compared individuals with different inter-ethnic interaction experiences on the basis of this metric. In order to make a more direct comparison with the theortical model described above, in the present analysis I use a single norm to calculate an absolute measure of cross-cultural competence. Using the classification strategy in Table A.1, I compare the degree and form of cross-cultural competence associated with different inter-ethnic experience types.

B.1.4. *A Norm of Fairness.* The present analysis focuses on one of the fourteen norms investigated in the previous study, which has very different distributions in the two ethnic groups, and is thus suited to the above metric of absolute cross-cultural competence. This norm for the fair division of an inheritance corresponds to Question 9 in Bunce (2020) and was presented to each interviewee as the following vignette:

(English translation)
An old woman has two new pots and two adult daughters. One daughter has her own two pots, but wants her mother's pots. The other daughter has no pots, and also wants her mother's pots. When the mother dies, who should inherit the pots? (Illustrated with a diagram.

Possible responses: one pot to each daughter; both pots to the daughter who has none)

(Spanish, as presented to Mestizo interviewees)
Hay una mujer vieja con dos ollas nuevas. Tiene dos hijas adultas. Una hija tiene sus propias dos ollas, pero quiere las ollas de su mamá. La otra hija no tiene ollas. También quiere las ollas de su mamá. Cuando la mamá se muere, ¿a quién debería heredarle las dos ollas?

(Matsigenka, as presented to Matsigenka interviewees)
Ogari tsinane okamake. Aityo pitieti ojiromanga otierira. Ainho piteni oshinto antaroni. Paniro oshinto aityo pitieti ojiromangane ashi iroro. Okogake oka otierira jiromanga. Ogari apiteni oshinto, mameri ojiromangane. Ariompa okogake oka otierira jiromanga. Tyani gakerone otierira jiromanga ashi iniro?

The question is intended to illustrate a norm for fair division according to right versus according to need. A response of "Both pots to the daughter who has none" was coded as positive, though this implies no judgement on my part as to which of the possible answers is "correct".

This question was inspired by my life history interviews with Mestizos, in which several people recounted instances of tension between siblings over the division of wealth belonging to a recently-deceased parent. In addition to norms of inheritance, this question is designed to investigate norms of fairness, e.g., division according to entitlement (equal shares to both daughters) or division according to need (both shares to the daughter who has less). My impression was that Mestizos tended to emphasize entitlement, and Matsigenka tended to emphasize need. Thus I hypothesized that Mestizos would give the negative response ("One pot to each daughter") and Matsigenka would give the positive response ("Both pots to the daughter who has none"). The qualification that the mother's pots are new (i.e., unused) when she dies is to circumvent the Matsigenka-typical norm of destroying or burying the used belongings of the deceased in order to avoid attracting a dangerous dead spirit. This is based on my observations and informal conversations during participation in a Matsigenka funeral, and corresponds with the interpretation of Shepard (2002).

A large majority (75%) of Matsigenka responded "Both pots to the daughter who has none", while most (68%) Mestizos responded "One pot to each daughter." An example of a typical Matsigenka explanation for giving both pots to the daughter who has none is because, "the other one already has pots." An example of a typical Mestizo response is, "I would give one [pot] to each one [i.e., daughter], because the first two pots [of the daughter who already has two] are her own. But I as a mother want to give the inheritance. And as I have two pots, and my other daughter has none, then I should give to each one [i.e., daughter] so that they don't fight. But I can't say to the other daughter [who already has two pots], 'Give her [i.e., your sister] the two pots', because [she] bought [her own two pots] with her own money. So I just give one [to each] in order to avoid [fights]. And with the disadvantage that one [daughter] has three [pots] and the other one."

B.1.5. *Inter-ethnic Interaction Experience Predictors.* The following measures of inter-ethnic interaction experience were recorded for each interviewee, and included in the statistical models as predictors. This is a subset of the predictors measured and used in the full analysis of Bunce (2020).

Education Experience (Edu = 1 : attended school with Mestizos)
All Mestizos attended primary and/or secondary school with other Mestizos, so all were coded as 1. Several Matsigenka interviewees grew up outside of Tayakome and went to either a boarding- or non-boarding primary school with Mestizos. These individuals were coded as 1. Most Matsigenka in Tayakome attended primary school in Tayakome, with Matsigenka teachers and all Matsigenka students. If this was an interviewee's only education experience, she or he was coded as 0. There is no secondary school in Tayakome. A few Matsigenka from Tayakome attended boarding secondary schools with Mestizos outside of Tayakome for at least four of the requisite five years, and some had additional educational training after high school (e.g., for tour-guide certification). These boarding school attendees were coded as 1. Two Matsigenka interviewees attended a boarding secondary school for a few months before either being expelled or leaving because they did not like it. These interviewees were coded as 0. The average amount of inter-ethnic education experience among Matsigenka scored as 1 was approximately 6.5 years. The number of Matsigenka interviewees coded as 1 who provided answers, respectively, to the preferred norm, in-group, and out-group guess questions: 17, 10, 10.

Wage Labor Employer Experience (Emp = 1 : experience employing Matsigenka)
Only one Matsigenka was coded as 1, as all other Matsigenka interviewees had never officially hired another Matsigenka as a wage laborer. 56 Mestizos (67%) were coded as 1 because they had, at some point,

paid money to a Matsigenka in return for labor. Most of this labor was short-term, on the order of one, or a few, days (e.g., harvesting a plantain field). However, 15 (27%) of the Mestizos scored as 1 employed Matsigenka for at least several months at a time (e.g., as crew for tour boats during tourist seasons). Mestizo interviewees coded as 1 (preferred norm, in-group guess, out-group guess): 56, 34, 34.

B.1.6. *Statistical Models.* In the present analysis, I focus on the probabilities that Matsigenka and Mestizos personally preferred a given norm for the fair division of inheritance (ego response), as well as the probabilities that they guessed that most members of their respective in-group and out-group preferred that norm (in-group response and out-group response, respectively). The response $y_{jt} = 0$ or 1 of individual j to target t (= ego, in-group, or out-group) is modeled as a logistic regression:

$$y_{jt} \sim \text{Binomial}(1, p_{jt}) \tag{A.1}$$

$$p_{jt} = \text{logit}^{-1}(\alpha_{jt}) \tag{A.2}$$

$$\alpha_{jt} = b_{jt} + mEDU_{t,\text{ETH}[j]} \cdot EDU_j \tag{A.3}$$

$$\begin{bmatrix} b_{t=\text{ego}} \\ b_{t=\text{in}} \\ b_{t=\text{out}} \end{bmatrix}_j \sim \text{MVNormal}\left(\begin{bmatrix} \mu_{b_{t=\text{ego}}} \\ \mu_{b_{t=\text{in}}} \\ \mu_{b_{t=\text{out}}} \end{bmatrix}_{\text{ETH}[j]}, \mathbf{S}_{\text{ETH}[j]} \right) \tag{A.4}$$

$$\mathbf{S}_{\text{ETH}[j]} = \begin{bmatrix} \sigma_{b_{t=\text{e}}} & 0 & 0 \\ 0 & \sigma_{b_{t=\text{i}}} & 0 \\ 0 & 0 & \sigma_{b_{t=\text{o}}} \end{bmatrix}_{\text{ETH}[j]} \mathbf{R}_{\text{ETH}[j]} \begin{bmatrix} \sigma_{b_{t=\text{e}}} & 0 & 0 \\ 0 & \sigma_{b_{t=\text{i}}} & 0 \\ 0 & 0 & \sigma_{b_{t=\text{o}}} \end{bmatrix}_{\text{ETH}[j]} \tag{A.5}$$

$$\begin{aligned} (mEDU_{t=\text{e}}, mEDU_{t=\text{i}}, mEDU_{t=\text{o}})_{\text{ETH=Mat}} &\sim \text{Normal}(0, 10) \\ (mEDU_{t=\text{e}}, mEDU_{t=\text{i}}, mEDU_{t=\text{o}})_{\text{ETH=Mes}} &\sim \text{Normal}(0, 10) \end{aligned} \tag{A.6}$$

$$\begin{aligned} (\mu_{b_{t=\text{e}}}, \mu_{b_{t=\text{i}}}, \mu_{b_{t=\text{o}}})_{\text{ETH=Mat}} &\sim \text{Normal}(0, 10) \\ (\mu_{b_{t=\text{e}}}, \mu_{b_{t=\text{i}}}, \mu_{b_{t=\text{o}}})_{\text{ETH=Mes}} &\sim \text{Normal}(0, 10) \end{aligned} \tag{A.7}$$

$$\begin{aligned} (\sigma_{b_{t=\text{e}}}, \sigma_{b_{t=\text{i}}}, \sigma_{b_{t=\text{o}}})_{\text{ETH=Mat}} &\sim \text{Exponential}(2) \\ (\sigma_{b_{t=\text{e}}}, \sigma_{b_{t=\text{i}}}, \sigma_{b_{t=\text{o}}})_{\text{ETH=Mes}} &\sim \text{Exponential}(2) \end{aligned} \tag{A.8}$$

$$(\mathbf{R}_{\text{ETH=Mat}}, \mathbf{R}_{\text{ETH=Mes}}) \sim \text{LKJcorr}(4) \tag{A.9}$$

The subscript ETH[j] is an indicator for the ethnicity (Matsigenka or Mestizo) of individual j. Estimating a separate mean intercept for each ethnicity has the same effect as including a main effect predictor for ethnicity in the linear model for α. I also allow the effects of inter-ethnic experience predictors (e.g., inter-ethnic education experience) to vary by ethnicity. This has the same effect as including an interaction of the predictors for ethnicity and inter-ethnic experience in the linear model for α. Thus, $mEDU_{t,\text{ETH}[j]}$ is the predictor for inter-ethnic education experience for target t for the ethnic group to which individual j belongs. It is multiplied by a binary indicator of individual j's education experience (0 or 1).

I had an *a priori* hypothesis that peoples' answers with respect to each of the targets (ego, in-group, and out-group) would covary, either positively or negatively. For instance, an individual's personally preferred norms may coincide with the norms she believes are held by the majority of her co-ethnics, and diverge from the norms she believes are preferred by most members of the out-group. I therefore model covariance among individuals' responses for each target by sampling each individual's three intercepts from a multivariate normal distribution. Exponential priors on standard deviations control ceiling and floor effects common in logistic models (McElreath 2016, pg 363-364).

After fitting the model, the variance-covariance matrix **S** contains estimates of residual covariance among individual-specific (i.e., random) intercepts (b_{jt}) across targets, after accounting for the variance in location among individuals for each target explained by their inter-ethnic experience. For instance, how individuals answered the vignette questions (ego responses) may covary with how they guessed members of the out-group answered the question (out-group guesses), even after accounting for the fact that certain types of inter-ethnic experience may affect both their own answers and their guesses about out-group individuals' answers. If true, estimates of the ego - out-group covariance contained in **S** will be non-zero.

Parameter estimation for each model was accomplished with RStan 2.17.3 (Stan Development Team 2018), running four Hamiltonian Monte Carlo chains in parallel until convergence was suggested by a high effective number of samples (> 500) and $\hat{R}$ estimates of 1.00 (McElreath 2016, pg 257). This entailed 3000 samples per chain, half of which were warm-up. In practice, a non-centered parameterization of the above model with Cholesky factorization of the correlation matrix **R** was fit in RStan (Stan Development Team 2017, pg 151). Data and statistical analysis scripts in R (R Core Team 2017) implementing RStan are available from Github at `https://github.com/jabunce/Bunce-2020-xcultural-competence`.

In addition to the above analysis, I also modeled individual ego responses and in- and out-group guesses to the fairness norm vignette question by leveraging individuals' responses and guesses to the other 13 norm questions investigated by Bunce (2020) using the item response theory (IRT) models developed as part of that study. This approach takes advantage of the fact that an individual's response to one question covaries with her responses to other questions. Thus, accounting for how she answered all 14 questions provides a more accurate posterior estimate for how she answered any one particular question. This analysis employs models m4 (incorporating only ethnicity), m5 (incorporating ethnicity and the Matsigenka predictor for inter-ethnic education), and m14 (incorporating ethnicity and the Mestizo predictor for inter-ethnic employer experience) defined in Tables S1 and S2 of Bunce (2020), resulting in posterior estimates analogous to those of the logistic regression described above. Note that one additional Matsigenka participant is included in this analysis, who had been excluded from Bunce (2020)'s analysis. That analysis included only individuals with complete information about additional domains of inter-ethnic interaction not considered here. Data and statistical analysis scripts in R (R Core Team 2017) implementing RStan are available from Github at `https://github.com/jabunce/Bunce-2020-xcultural-competence`.

B.1.7. *Phenotype Frequencies from Posteriors.* The posterior distribution of the probability that an average (randomly chosen) Matsigenka would give the positive response to the fairness norm preference question would be expected to have most of its probability density $> \frac{1}{2}$, as 75% of Matsigenka interviewees gave the positive response. Similarly, the corresponding posterior distribution for Mestizos would be expected to have most of its probability density $< \frac{1}{2}$, as only 32% of Mestizos gave the positive response. However, due to uncertainty in the model estimates, some probability density for each of these posteriors may be found on the side of $\frac{1}{2}$ opposite of that expected, and this must be taken into account when computing model-estimated probabilities (frequencies) of the various phenotypes in these two ethnic groups. As noted above in section B.1.1 Defining and Measuring CCC, the metric developed here is designed for situations when such unanticipated posterior density is very low. This is indeed the case here, as can be seen in Appendix Figure A.1.

Calculation of the expected phenotype frequencies from model posteriors is accomplished as follows. Ignoring inter-ethnic interaction experience, the posterior estimate of the probability that an average (randomly chosen) Matsigenka would prefer the norm coded as positive is

$$M_{\mathrm{e}} = p_{\mathrm{matsi,ego}} = \mu_{b_{\mathrm{matsi,ego}}} \tag{A.10}$$

Compare Equation A.10 with Equations A.2-A.4. $M_{\mathrm{e}}[y]$ is the y^{th} sample of this probability in the posterior sample set $[1, Y]$. For convenience, we can equate the Matsigenka ethnic group with group S in the theoretical model, and a positive response to the vignette question with norm 1 in the theoretical model. Therefore, for posterior sample y, the estimated mean probability (frequency) of the uni- and cross-culturally competent phenotypes $S11$ and $S1X$ among Matsigenka is:

$$p_{S11}[y] = M_e[y] \cdot M_i[y] \cdot M_o[y] \tag{A.11}$$

$$p_{S1X}[y] = \begin{cases} M_e[y] \cdot M_i[y] \cdot (1 - M_o[y]); & M_e[y] > \frac{1}{2} \ \& \ Z_e[y] < \frac{1}{2} \\ M_e[y] \cdot (1 - M_i[y]) \cdot M_o[y]; & M_e[y] < \frac{1}{2} \ \& \ Z_e[y] > \frac{1}{2} \\ M_e[y] \cdot (1 - M_i[y]) \cdot (1 - M_o[y]); & M_e[y] < \frac{1}{2} \ \& \ Z_e[y] < \frac{1}{2} \\ 0; & M_e[y] > \frac{1}{2} \ \& \ Z_e[y] > \frac{1}{2} \end{cases} \tag{A.12}$$

where M_i and M_o are the posterior estimates of the probabilities that an average Matsigenka would guess that an average Matsigenka (in-group) or Mestizo (out-group), respectively, prefers the positive norm. Z_e is the posterior estimate of the probability that an average Mestizo prefers the positive norm. These posterior probabilities are calculated using a procedure analogous to Equation A.10. The four cases of Equation A.12 depend on the estimated probabilities of the preferred (ego) norm among Matsigenka and Mestizos (M_e and Z_e, respectively). For the reasonable application of this metric, the first case should heavily dominate all others, which it does here (Appendix Figure A.1).

Note that, in the last case of Equation A.12, the cross-culturally competent phenotype $S1X$ would be calculated as $M_e[y] \cdot M_i[y] \cdot M_o[y]$, and would thus be equivalent to the uni-culturally competent phenotype $S11$ according to the above definitions. When most people in both the in-group and the out-group prefer the same norm that you do, cross-cultural competence is achieved by knowing only that norm. Here, the (arbitrary) decision is made to classify as uni-culturally competent all phenotypes that prefer and know only one norm (Equation A.11), even if they could also be classified in other ways.

The first case in Equation A.12 corresponds to the phenotype classification in Table A.1. Frequencies of the other Matsigenka phenotypes, $S22$ and $S2X$, as well as the other four phenotype classifications in Table A.1 and the equivalent Mestizo phenotypes (equated with group L in the theoretical model) are calculated analogously.

B.1.8. *Empirical Results.* Appendix Figure A.1 plots the posterior estimates of the mean probability of preferring norm 1 (ego response = 1 rather than 0) for Matsigenka and Mestizos. Note that nearly all of the posterior probability mass for Matsigenka falls above 0.5, and for Mestizos falls below 0.5, corresponding with the first case in Equation A.12. This pattern is even more clearly apparent in Appendix Figure A.2, derived from the corresponding IRT model m4 in Bunce (2020), which leverages individuals' answers to 14 norm vignette questions in order to estimate individuals' probability of preferring norm 1, of interest here (question 9 in Bunce 2020). This supports the assumption that most Matsigenka personally prefer norm 1, while most Mestizos personally prefer norm 2 (ego response = 0).

Appendix Figures A.3 and A.4 plot the distributions of the posterior estimates for the mean probabilities (frequencies) of each phenotype in the Matsigenka and Mestizo ethnic groups (e.g., $p_{S1X}[y]$, for all $y \in [1, Y]$ posterior samples) at one point in time, namely, the year 2013 when data were collected. Note the similarities with the same estimates derived from the IRT models in Bunce (2020), shown in Appendix Figures A.5 and A.6.

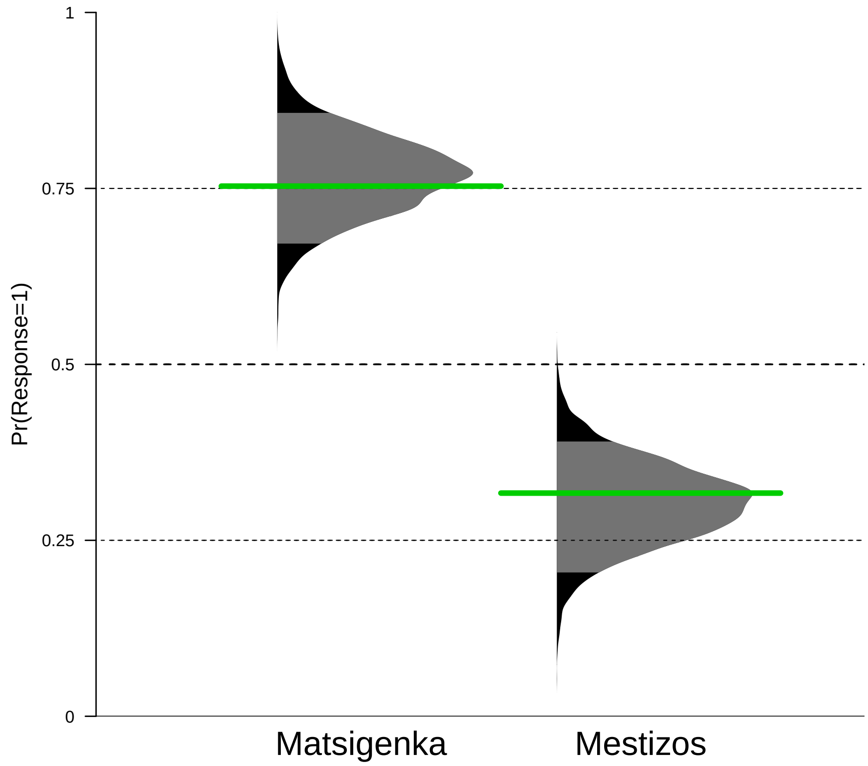

FIGURE A.1. Posterior estimates of the mean probability of preferring norm 1 (giving the positive ego response to the vignette question about fairness) for Matsigenka and Mestizos, derived from the logistic regression in Equations A.1-A.9, with only a random intercept (no inter-ethnic experience predictors). 90% highest posterior density intervals are shown in grey. Green lines indicate raw proportions of the 76 Matsigenka and 82 Mestizo participants who gave the positive response.

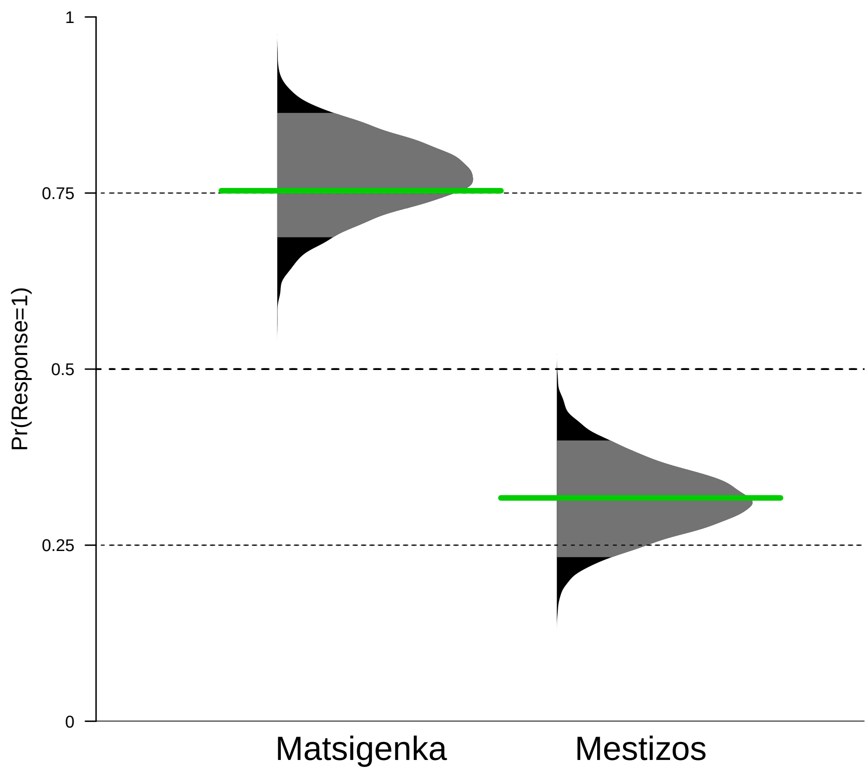

FIGURE A.2. Posterior estimates of the mean probability of preferring norm 1 (giving the positive ego response to question 9) for Matsigenka and Mestizos, derived from IRT model m4 in Bunce (2020). 90% highest posterior density intervals are shown in grey. Green lines indicate raw proportions of the 76 Matsigenka and 82 Mestizo particpants who gave the positive response.

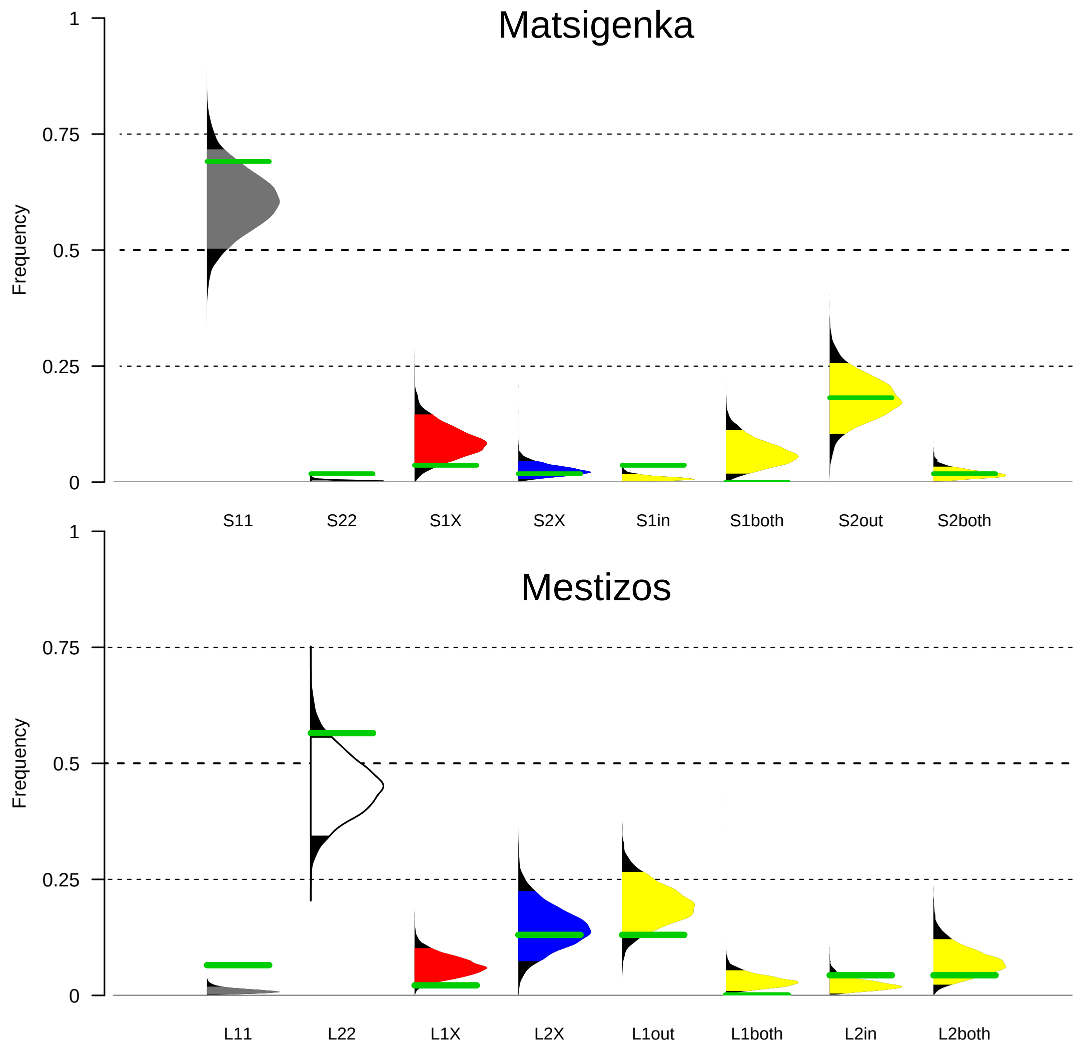

FIGURE A.3. Posterior estimates of the mean probabilities (frequencies) of all possible phenotypes among minority Matsigenka (group S, top plot) and majority Mestizos (group L, bottom plot), with respect to norms of fairness, derived from the logistic regression in Equations A.1-A.9, with only a random intercept (no inter-ethnic experience predictors). 90% HPDI are shown in grey. Green lines indicate raw proportions of the 57 Matsigenka and 46 Mestizo particpants classified as each phenotype.

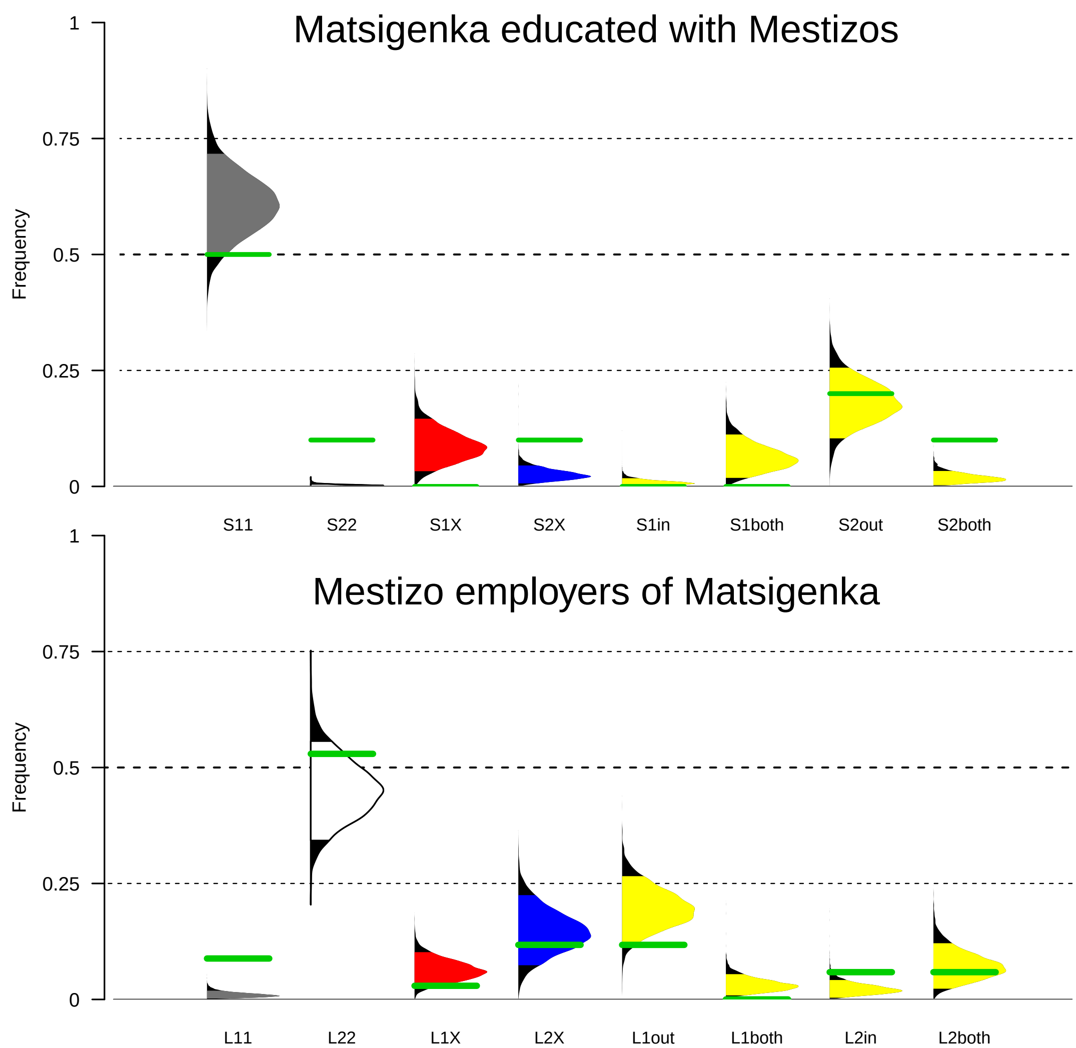

FIGURE A.4. Posterior estimates of the mean probabilities (frequencies) of all possible phenotypes among Matsigenka educated among Mestizos and Mestizo employers of Matsigenka, with respect to norms of fairness, derived from the logistic regression in Equations A.1-A.9, with corresponding inter-ethnic experience predictors. 90% HPDI are shown in grey. Green lines indicate raw proportions of the 10 Matsigenka and 34 Mestizo participants with the relevant inter-ethnic experience classified as each phenotype.

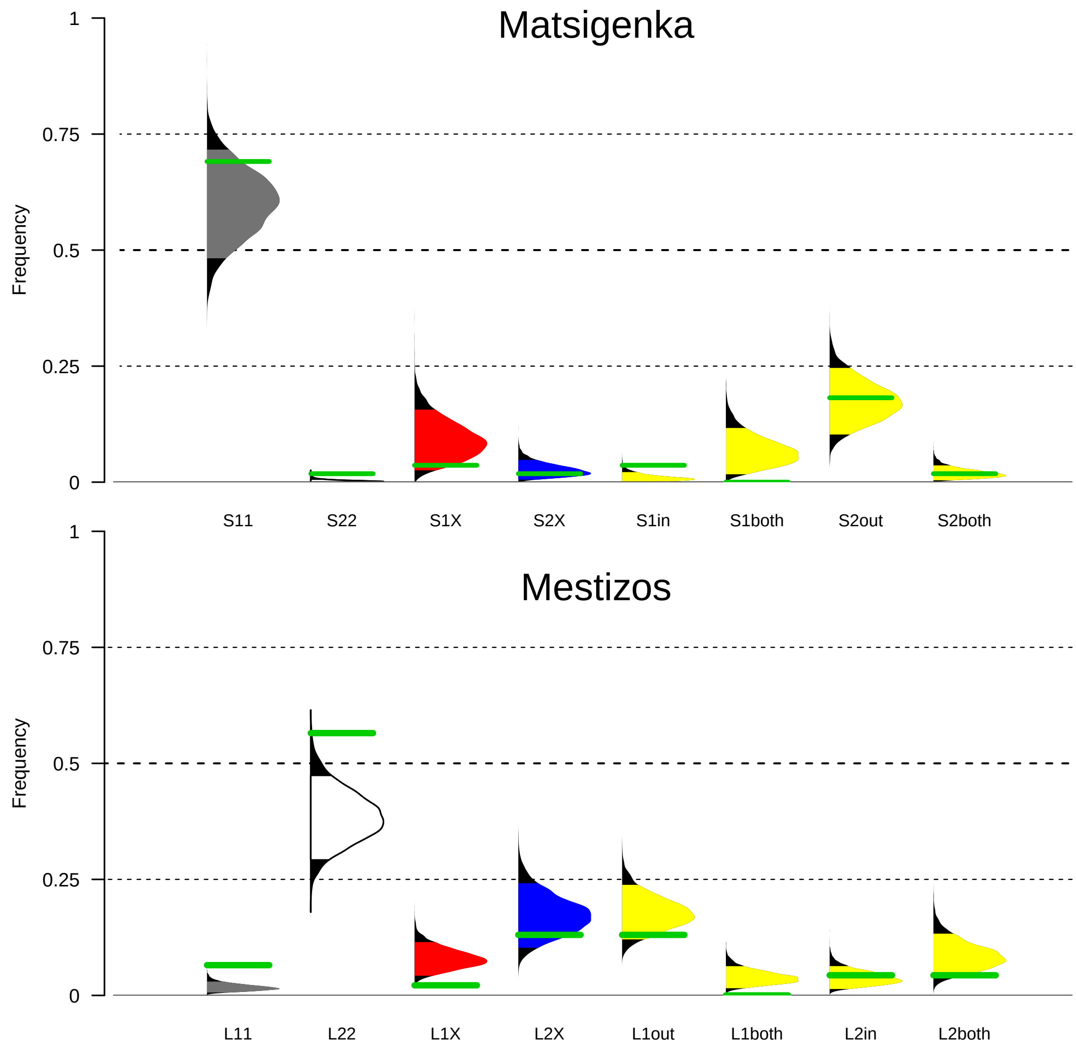

FIGURE A.5. Posterior estimates of the mean probabilities (frequencies) of all possible phenotypes among minority Matsigenka (group S, top plot) and majority Mestizos (group L, bottom plot), with respect to norms of fairness, derived from IRT model m4 in Bunce (2020). 90% HPDI are shown in grey. Green lines indicate raw proportions of the 57 Matsigenka and 46 Mestizo particpants classified as each phenotype.

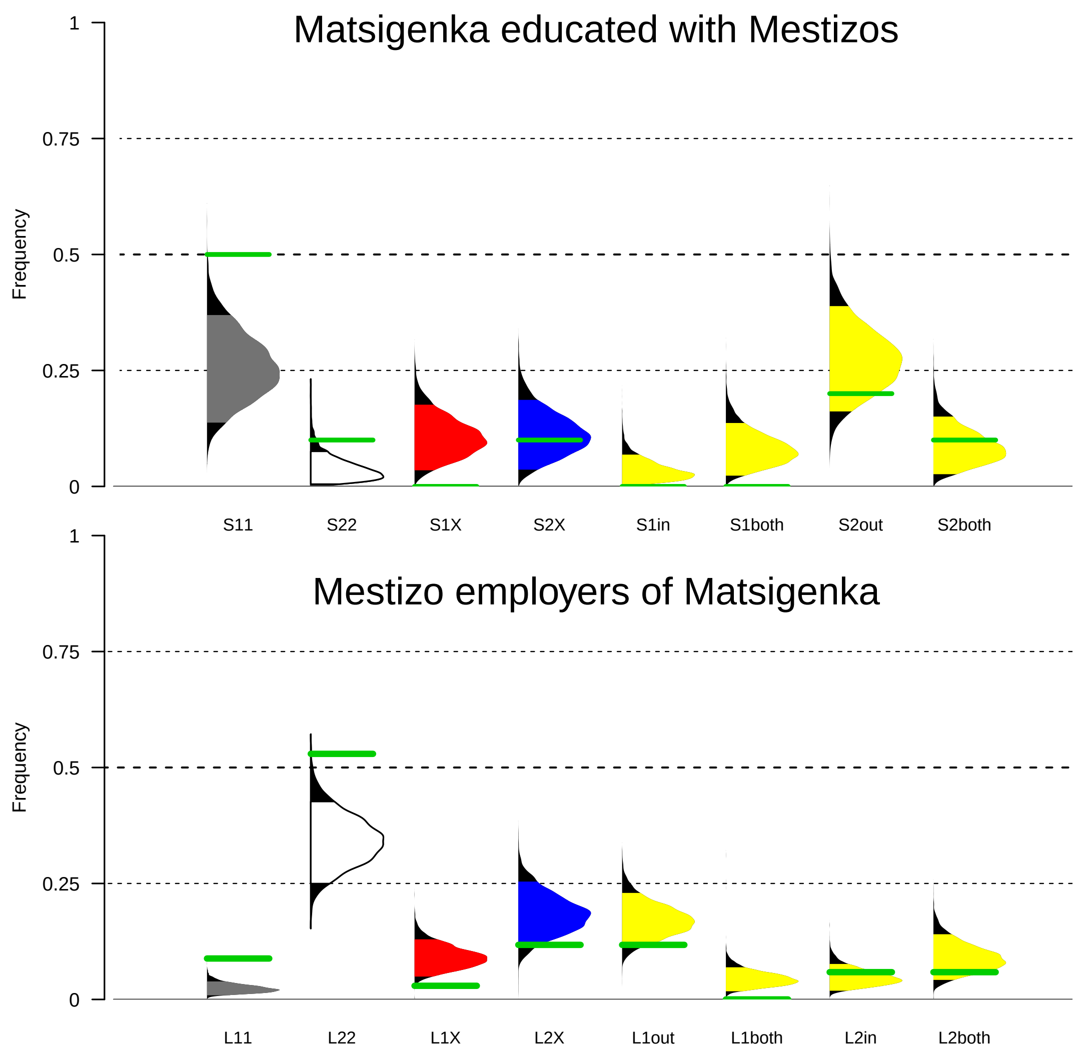

FIGURE A.6. Posterior estimates of the mean probabilities (frequencies) of all possible phenotypes among Matsigenka educated among Mestizos (top plot, IRT model m5 in Bunce 2020) and Mestizo employers of Matsigenka (bottom plot, IRT model m14), with respect to norms of fairness. 90% HPDI are shown in grey. Green lines indicate raw proportions of the 10 Matsigenka and 34 Mestizo participants with the relevant inter-ethnic experience classified as each phenotype.

For comparison with the theoretical model, Figure 1A in the main text shows posterior estimates for the frequencies of only the four UCC and CCC phenotypes in Table A.1, standardized to sum to one. These estimates are derived from the IRT models in Bunce (2020). The top row of Figure 1A uses the posteriors of IRT model m4 to compare mean estimated phenotype frequencies of all Matsigenka and all Mestizos, regardless of inter-ethnic experience. Comparing the top row of Figure 1A with the top row of Figure 1B, it can be seen that estimated phenotype frequencies among Matsigenka and Mestizos roughly correspond with those predicted by the theoretical model in early stages of the dynamics (i.e., near the 5th time step) when valuation of in-group identity (i) is lower than it is in the top row of Figure 2A.

The bottom left and right plots in Figure 1A use the posteriors of IRT models m5 and m14, respectively, to compare mean estimated phenotype frequencies among the subset of Matsigenka with education experience among Mestizos, and the subset of Mestizos who have experience hiring Matsigenka laborers. (See Tables S1 and S2 in Bunce 2020 for model definitions). Based on a previous analysis (Bunce 2020), Matsigenka who were educated in Mestizo schools are expected to have experienced an intensive period of inter-ethnic interaction, which facilitated the adoption of many Mestizo-typical norms. These individuals are therefore expected to be more cross-culturally competent than their fellow Matsigenka. Similarly, Mestizos who have experience as employers of Matsigenka are expected to have learned many Matsigenka-typical norms (though they may not personally prefer these norms), and are therefore expected to be most cross-culturally competent. Comparing the bottom row of Figure 1A with the bottom row of Figure 1B (where in-group affinity a is low), it can be seen that estimated phenotype frequencies among Matsigenka and Mestizos with such inter-ethnic interaction experience roughly correspond with those predicted by the theoretical model in early stages of the dynamics (e.g., near the 5th time step), given a high degree of inter-ethnic interaction. Thus, this analysis shows that, despite its simplistic assumptions, the theoretical model of the dynamics of cross-culturally competent phenotypes may, under certain conditions, plausibly approximate such dynamics in real populations, as measured using field-interview methods.

B.2. **Theoretical Model Design.** An arbitrarily large population is divided into a smaller group S and a larger group L (relative sizes defined below). Each person in either group prefers one of two norms: 1 or 2. A proportion p_{S11} of group S comprises individuals who prefer norm 1 and always employ norm 1 whenever they interact with other people. p_{S22} is the analogous proportion of individuals who prefer norm 2. $S11$ and $S22$ are "uni-culturally competent" (UCC) because they only interact using their preferred norm. A proportion p_{S1X} of group S comprises individuals who prefer norm 1, but who are capable of employing either norm 1 or norm 2 during interactions. Thus, when interacting with $S11$ or $S22$ individuals, $S1X$ employs norm 1 or norm 2, respectively. $S1X$ individuals are "cross-culturally competent" (CCC). Analogously, p_{S2X} is the proportion of CCC individuals in group S who prefer norm 2. p_{L11}, p_{L22}, p_{L1X}, and p_{L2X} are the analogous proportions of UCC and CCC individuals in group L. Initially, group S comprises a large majority of individuals who prefer norm 1 (i.e., large p_{S11} and/or p_{S1X}), and group L comprises a large majority that prefers norm 2 (i.e., large p_{L22} and/or p_{L2X}). Thus, the two groups represent ethnic groups characterized by distinctive distributions of cultural norms.

B.2.1. *Interaction Phase.* The model has two phases: interaction and updating. In the interaction phase, pairs of individuals play a coordination game. If both individuals can employ the same norm, both receive a base coordination payoff of 1. If they cannot, both receive a payoff of 0. The parameter $a \in [0, 1]$ measures the strength of in-group affinity during interaction, such that a member of group S interacts with a randomly-chosen in-group member with probability a_S, and with a randomly-chosen out-group member with probability $1 - a_S$. a_L is the analogous probability for members of group L. The parameter a is set exogenously in the model, and does not evolve (see below). Under the assumption that all individuals engage in exactly one interaction per interaction phase, $\frac{1-a_L}{1-a_S}$ is the ratio of the population sizes of group S to group L (inspired by Bruner 2019 and Mohseni et al. 2019). To represent the context of minority/majority ethnic groups, this ratio is fixed at 1/2 in all model simulations, such that group L is twice the size of group S (i.e., an L individual is half as likely as an S individual to interact with an out-group member). For example, assume that group S members have a probability $1 - a_S = 0.4$ of interacting with members of group L. Thus, on average, 40% of S members interact with the out-group. If group L is larger than group S, the number of L members who interact with S members will constitute $< 40\%$ of group L – say, for example, 20% (such that $1 - a_L = 0.2$). Thus, 20% of the size of group L is equal to 40% of the size of group S. Group L is twice the size of group S and $\frac{1-a_L}{1-a_S} = \frac{0.2}{0.4} = \frac{1}{2}$.

When a cross-culturally competent individual interacts with a uniculturally competent individual who prefers a different norm (e.g., $S1X$ and $L22$), the cross-culturally competent individual will employ, and successfully coordinate using, a norm different from the one that she prefers. In such situations, the cross-culturally competent individual suffers a cost $c \in [0, 1]$, which is the difficulty (e.g., cognitive dissonance: Festinger 1962) associated with performing an action that conflicts with one's personally preferred norm. For simplicity, c is assumed to be constant for all cross-cultural phenotypes. c has an upper bound at the base coordination payoff of 1, such that cross-culturally competent individuals never receive a payoff for successfully coordinating that is lower than the payoff for unsuccessfully coordinating, i.e., 0. When two cross-culturally competent individuals who prefer different norms interact (e.g., $S1X$ and $S2X$), they always coordinate, but the norm they use to do so is chosen at random. Thus, on average, for half of such interactions each individual suffers the cognitive dissonance cost c. Note that two CCC individuals from the same group who both prefer a norm that differs from the norm perceived to be most common in their in-group will coordinate using the norm that they both prefer, rather than the norm typical of their fellow in-group members. Thus, this model does not implement "preference falsification" (Kuran 1995) on the part of CCC individuals when this is unnecessary to ensure coordination. Independent of payoffs to coordination interactions, there is a general learning cost, $m \geq 0$, associated with expending effort to learn a non-preferred norm and thereby become cross-culturally competent. It is assumed that this initial learning cost is payed out over the course of one's life (i.e., in each interaction step). Conceiving of m as such an opportunity cost seems natural, as the time an individual invests to learn a new norm is time that she cannot invest to learn a different skill that will serve her over the course of her life.

It is assumed that individuals cannot see each others' phenotypes, and thus, they cannot assort by phenotype or by some marker that covaries with phenotype. This also means that cross-culturally competent individuals do not initially know which norm to use when they are paired with an interaction partner. I assume that at the beginning of each inter-individual interaction, a sequence of norm proposals is enacted. This sequence is not explicitly modeled. Each individual simultaneously makes a proposal of the norm that she would like to use to coordinate. This proposed norm is the personally preferred norm of each individual. A cross-culturally competent individual, upon seeing her partner's proposal, changes her proposal to match her partner's if their intial proposals do not coincide. If this partner is another cross-culturally competent individual, the member of the interaction pair who changes her proposal first is chosen at random. A uni-culturally competent individual cannot change her proposal. After this sequence, the proposed norms of the two partners either match (if at least one is cross-culturally competent) or do not match (if both are uni-culturally competent and prefer different norms). The (mis)coordination interaction then proceeds.

Coordination interactions between individuals from different ethnic groups often entail imbalances in payoffs accruing to the participants. In other words, one of the two members of an inter-ethnic interaction pair often receives a higher subjective benefit from successful coordination than does the other. Following previous work (Bunce and McElreath 2017, 2018), the individual who receives the higher coordination payoff has lower bargaining power in that interaction (see also O'Conner 2019). Let $b_S \geq 0$ and $b_L \geq 0$ be the extra coordination benefits accruing to individuals from group S and group L, respectively, who successfully coordinate with individuals from the other ethnic group. To represent a context of structural inequality, such that members of the minority group S have lower bargaining power during interactions with members of the majority group L, it is assumed that $b_S > b_L$ in all subsequent analyses.

Note that this implementation of inter-group coordination payoffs (b) requires additional assumptions. For instance, for any $b > 0$, provided inter-group coordination can be achieved, individuals from one group appear to have an incentive to coordinate only with out-group members, and not at all with in-group members. Thus, were in-group interaction affinity (a) endogenous to the model, we might expect it to evolve to 0. A constant exogenous a, as employed in this model, is applicable under the assumption that inter-group interaction is exogenously restricted to a certain level, such as by limited available transportation between geographically separated communities.

Alternatively, one could assume complementarities between in-group and out-group coordination payoffs, such that both types of payoff together are required to maximize utility. For instance, if in-group coordination yields payoffs of food and out-group coordination yields payoffs of luxury goods, coordinating with members of only one of the two groups would result in sub-optimal utility, as individuals want both food and luxury goods. Furthermore, luxury goods become valuable only after the point where one has obtained sufficient food. Figure A.7 represents a context in which all inter-individual interactions result in coordination (i.e., all individuals have the same norms and/or are cross-culturally competent). Thus, the probability of in-group affinity (a) is equal to the probability of obtaining a payoff of 1. Alternatively, a can be thought of as the quantity of payoff obtained for a given (constant) number of in-group interactions per time step. Payoffs exhibit diminishing returns, such that each individual payoff (increment of a) decreases in subjective value as more payoffs are obtained (Figure A.7:

val_i is decreasing in a). Additionally, in-group payoffs exhibit a satiation limit a^*, such that payoffs obtained in excess of this limit are not valued. Only once the in-group satiation limit has been reached, are payoffs from out-group interactions valued. Out-group payoffs also exhibit diminishing returns in subjective value (val_o is decreasing in $1-a$, the probability of out-group affinity). At a^*, the value of an out-group coordination payoff is $1 + b^*$. As an individual becomes less likely to obtain out-group payoffs (as a increases), she values each out-group payoff more than $1 + b^*$.

Let utility from in-group coordination, util_i, represent the sum of the subjective value obtained from in-group payoffs. This is calculated as the value of an in-group payoff (val_i), which is a function of the frequency with which it is obtained (a), times the frequency with which it is obtained. Utility from out-group coordination, util_o, is calculated analogously. Let total utility be the sum of util_i and util_o, and assume individuals attempt to maximize this. As shown in the lower plot of Figure A.7, total utility is maximized at an in-group affinity of a^*, at which point (from the upper plot) an out-group coordination payoff is subjectively valued at $1 + b^*$.

In the model in the main text, the parameters a and b are exogenously assigned, and therefore do not evolve. This model assumption corresponds to a context in which in-group and out-group coordination payoffs are valued in such a way that they are complementary. A given a represents the satiation limit for in-group payoffs, and the correspondingly chosen value of $1 + b$ represents the subjectively-valued out-group coordination payoff, conditional on a probability $1 - a$ of interacting with out-group individuals. As shown above, under such conditions, individuals have no incentive to attempt any more or any fewer interactions with the out-group than those represented by the exogenous value of a.

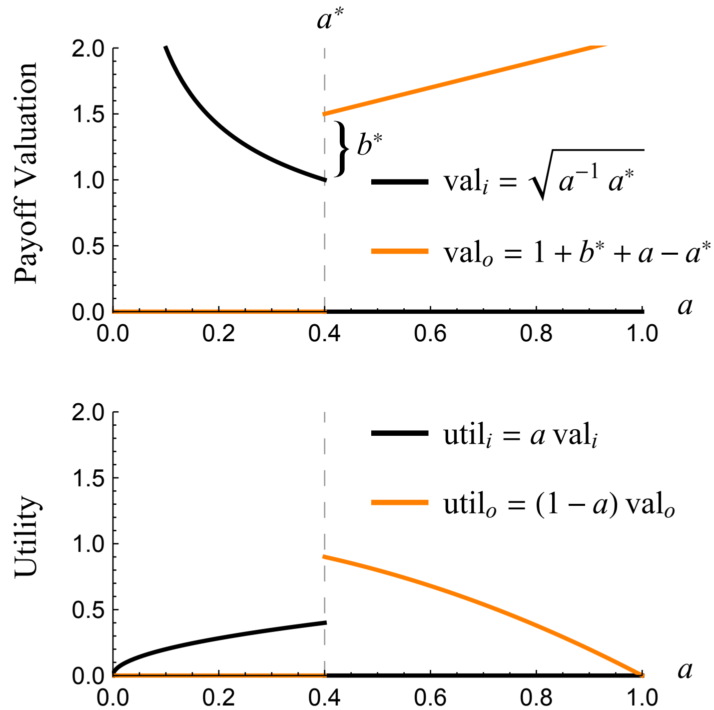

FIGURE A.7. Example of the relationship between in-group affinity a and inter-group coordination payoff $1+b$, when all inter-individual interaction results in coordination. Upper plot: The subjective valuation of a payoff to coordination with in-group members (val_i) as a function of a is given for the interval $a \in [0, a^* = 0.4]$, and is 0 otherwise. The valuation of a payoff to coordination with out-group members (val_o) is given for the interval $a \in [a^*, 1]$, and is 0 otherwise. Both valuation functions exhibit diminishing returns with, respectively, increasing interaction with in-group members (a) and with out-group members ($1 - a$). The satiation limit for in-group coordination payoffs is a^*, only above which are out-group payoffs valued. Lower plot: An individual's utility is the product of the value individuals place on a payoff and the probability of obtaining the payoff. On average, individuals are assumed to choose an in-group affinity level a^* that maximizes the sum of their utility arising from in-group interactions (util_i) and out-group interactions (util_o). The value $1 + b^* = 1.5$ is the average valuation of the out-group coordination payoff at this level of in-group affinity.

Payoffs from coordination interactions to uni-culturally competent and cross-culturally competent individuals are shown in Table A.2. Average payoffs for each phenotype in group S, incorporating coordination payoffs and costs of cognitive dissonance and learning are shown in Equations A.13-A.16.

TABLE A.2. Payoffs from coordination interactions to Person 1 from Group S[a]

		Person 2 norm			
Pairing[b]	Person 1 norm	Norm 1	Norm 2	Norm 1	Norm 2
U – U	Norm 1	1	0	$1 + b_S$	0
	Norm 2	0	1	0	$1 + b_S$
C – C	Norm 1	1	$1 - \frac{1}{2}c$	$1 + b_S$	$1 + b_S - \frac{1}{2}c$
	Norm 2	$1 - \frac{1}{2}c$	1	$1 + b_S - \frac{1}{2}c$	$1 + b_S$
C – U	Norm 1	1	$1 - c$	$1 + b_S$	$1 + b_S - c$
	Norm 2	$1 - c$	1	$1 + b_S - c$	$1 + b_S$
U – C	Norm 1	1	1	$1 + b_S$	$1 + b_S$
	Norm 2	1	1	$1 + b_S$	$1 + b_S$
	Person 2 is:	in Group S		in Group L	

[a]Payoffs to a member of Group L are identical after substituting b_L for b_S.

[b]U–U: two uni-culturally competent individuals interact. C–C: two cross-culturally competent individuals interact. C–U: Person 1 is cross-culturally competent, Person 2 is uni-culturally competent. U–C: Person 1 is uni-culturally competent, Person 2 is cross-culturally competent.

$$w_{S11} = a_S(p_{S11} + p_{S1X} + p_{S2X}) + \\ (1 - a_S)\left[p_{L11}(1 + b_S) + p_{L1X}(1 + b_S) + p_{L2X}(1 + b_S)\right] \quad (A.13)$$

$$w_{S1X} = a_S\left[p_{S11} + p_{S1X} + p_{S2X}(1 - \tfrac{1}{2}c) + p_{S22}(1 - c)\right] + \\ (1 - a_S)\left[p_{L11}(1 + b_S) + p_{L1X}(1 + b_S) + \\ p_{L2X}(1 + b_S - \tfrac{1}{2}c) + p_{L22}(1 + b_S - c)\right] - m \quad (A.14)$$

$$w_{S2X} = a_S\left[p_{S11}(1 - c) + p_{S1X}(1 - \tfrac{1}{2}c) + p_{S2X} + p_{S22}\right] + \\ (1 - a_S)\left[p_{L11}(1 + b_S - c) + p_{L1X}(1 + b_S - \tfrac{1}{2}c) + \\ p_{L2X}(1 + b_S) + p_{L22}(1 + b_S)\right] - m \quad (A.15)$$

$$w_{S22} = a_S(p_{S1X} + p_{S2X} + p_{S22}) + \\ (1 - a_S)\left[p_{L1X}(1 + b_S) + p_{L2X}(1 + b_S) + p_{L22}(1 + b_S)\right] \quad (A.16)$$

Payoffs to phenotypes in group L are found by reversing all group and norm indices in the subscripts. Parameters a_S, b_S, c, and m are defined above. Note that these payoff expressions are not directly incorporated into the model. Rather they are modified below in Equations A.19-A.22.

B.2.2. *Updating Phase.* Each individual is aware of whether, in the interaction phase, she either failed to coordinate (if uni-culturally competent) or suffered the cognitive dissonance cost c (if cross-culturally competent). If neither occurred, in the updating phase she retains her current phenotype. However, if either occurred, she evaluates whether it is worthwhile to change her phenotype, according to the set of rules shown in Figure A.8. This feature of the model operationalizes an assumption of psychological inertia: People tend to persist in their current beliefs and behavior until a non-optimal outcome inspires them to re-evaluate, and potentially modify, those beliefs or behavior (see Bicchieri 2006, pg 148-149).

The allowable phenotype transitions in Figure A.8 reflect the assumption that once an individual learns a norm, it cannot be unlearned. Thus, uni-culturally competent individuals can become cross-culturally competent. However, though cross-culturally competent individuals may change the norm that they prefer, they cannot forget a norm that they know and thereby become uni-culturally competent. Cross-cultural competence (of some kind), is thus an absorbing state

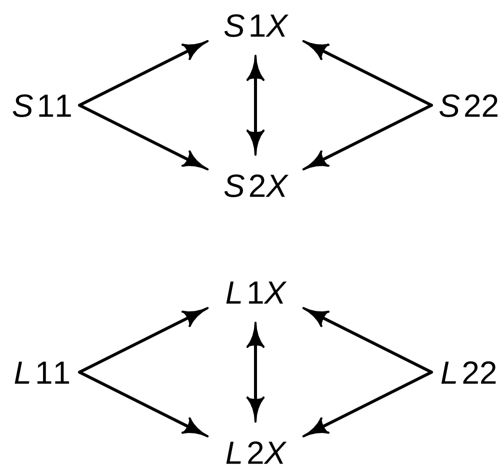

FIGURE A.8. Allowable phenotype transitions. Group S is the low-power minority group and L is the high-power majority group. The designations 11 and 22 refer to uni-culturally competent phenotypes preferring norm 1 or norm 2, respectively. Designations 1X and 2X refer to cross-culturally competent phenotypes preferring norm 1 or 2, respectively.

of the model. The theoretical question of interest is which form of cross-cultural competence ($S1X$ or $S2X$) evolves to high frequency in minority group S.

An individual who failed to coordinate, or who suffered a cognitive dissonance cost, in the interaction phase makes a strategically rational choice about whether to change her phenotype, and, if so, which new phenotype to adopt. However, she makes this choice based on imperfect information about the phenotypes of others. It is assumed that an individual cannot see another individual's preferred norm. Rather, one must infer an individual's preferred norm from one's observation of the norm that she attempts to use during a single interaction. When a cross-culturally competent individual interacts using a norm that she does not prefer, others will incorrectly infer that she prefers a norm that in fact she does not. This tendency for inaccurate inference of internal belief from observed behavior underlies the well-documented psychological phenomenon of pluralistic ignorance (Katz and Allport 1931; Prentice and Miller 1993). It is assumed that an individual contemplating a phenotype change attempts to determine which norms she is likely to encounter in the next interaction phase. She therefore surveys all (or a representative sample) of her fellow in-group members and observes: 1) which norms they attempted to use during the previous interaction phase when they were paired with other in-group members like her; and 2) which norms out-group members attempted to use when paired with her fellow in-group members. In small-scale societies, such information about co-ethnics' successful and unsuccessful social interactions may become public knowledge as a result of gossip (Gluckman 1963; Wiessner 2005). Based on this information, an individual infers (potentially inaccurately) the probability with which she is likely to be confronted with each norm in the next interaction phase if paired with either an in- or an out-group member. It is assumed that individuals, even those who are cross-culturally competent themselves, are unaware that other individuals in the population may be cross-culturally competent. Thus, a person assumes that the norms she observes fellow in-group members to use with other in-group members are these individuals' preferred norms, and that they would attempt to use them in any potential future interaction with her (analogous assumption for the norms observed to be used by out-group individuals). Equations A.17 and A.18 give, respectively, the perceived probability that an S member will attempt to use norm 1 and the probability that an L member will attempt to use norm 2, when paired with a member of group S in the next interaction phase. These are simply the probabilities of these norms being used with group S members in the previous interaction phase. For convenience in what follows, I refer to these probabilities, respectively, as the effective frequency of norm 1 in the in-group and the effective frequency of norm 2 in the out-group, as perceived by a member of group S.

$$\tilde{p}_{S1in} = p_{S11}(p_{S11} + p_{S1X} + p_{S2X} + p_{S22}) + p_{S1X}(p_{S11} + p_{S1X} + \tfrac{1}{2}p_{S2X}) + p_{S2X}(p_{S11} + \tfrac{1}{2}p_{S1X}) \quad \text{(A.17)}$$

$$\tilde{p}_{L2out} = p_{L22}(p_{S22} + p_{S2X} + p_{S1X} + p_{S11}) + p_{L2X}(p_{S22} + p_{S2X} + \tfrac{1}{2}p_{S1X}) + p_{L1X}(p_{S22} + \tfrac{1}{2}p_{S2X}) \quad \text{(A.18)}$$

The probabilities $\tilde{p}_{S2out}$ and $\tilde{p}_{L2in}$ are found by reversing all group and norm indices in the subscripts of Equations A.18 and A.17, respectively. An individual's decision about if and how to change her phenotype is based on the payoff that she expects to receive in the next interaction phase, given her knowledge of the probability of interacting with an in-group versus an out-group member (a), and her inferred probability of interacting with people who prefer a given norm in each of these groups. Thus, such an individual mentally compares the anticipated payoffs accruing, in the next interaction phase, to her current phenotype and to the phenotype(s) to which she can potentially change. However, as explained next, such anticipated payoffs may be modified by an effect of group identity.

Many norms of social coordination are associated with cultural group identity. Language is perhaps the most common of these (sets of) norms. Other norms often associated with cultural identity that may also be deployed in contexts of social coordination include norms of fairness and honor (Cohen et al. 1996), and norms of what is appropriate to eat (Baer 2004) and how (or if) to dress (Gow 1993). If an individual values her cultural identity, then the knowledge of, and ability to use, such a norm may entail a positive utility (Akerlof and Kranton 2000). Furthermore, this utility may be independent of the frequency of opportunities she has to actually use the norm for coordination. For example, a Chinese immigrant to the United States who values her Chinese cultural identity may derive utility from the fact that she knows Mandarin, even if she has no opportunities to speak it in her new home. I hypothesize that this identity-based utility depends both on the population-level distribution of the norm and on characteristics of individual psychology. In this model, identity-based utility increases with: 1) the perceived frequency of the norm within the in-group; 2) the perceived rarity of the norm within the out-group (i.e., the degree to which the norm distinguishes the in-group from the out-group); and 3) the degree to which an individual values the cultural identity of her in-group, operationalized as $i \geq 0$. For simplicity, i is assumed to be constant for all individuals, and individuals cannot punish others for behaving in ways perceived to be inconsistent with the group identity ascribed to them (unlike Ackerlof and Kranton 2000). Note that, though i is the same for all individuals, its effect on payoffs is mediated by perceived in- and out-group norm frequencies. Thus, the effect of i on payoffs is potentially different for members of each group.

Average anticipated payoffs for each phenotype of group S, incorporating inferred effective frequencies of preferred norms and identity-based utility are shown in Equations A.19-A.22.

$$\tilde{w}_{S11} = a_S\tilde{p}_{S1in} + (1 - a_S)(1 - \tilde{p}_{L2out})(1 + b_S) + i\tilde{p}_{S1in}\tilde{p}_{L2out} \quad \text{(A.19)}$$

$$\begin{aligned}\tilde{w}_{S1X} = {} & a_S\left[\tilde{p}_{S1in} + (1 - \tilde{p}_{S1in})(1 - c)\right] + \\ & (1 - a_S)\left[(1 - \tilde{p}_{L2out})(1 + b_S) + \tilde{p}_{L2out}(1 + b_S - c)\right] - \\ & m + i\tilde{p}_{S1in}\tilde{p}_{L2out}\end{aligned} \quad \text{(A.20)}$$

$$\begin{aligned}\tilde{w}_{S2X} = {} & a_S\left[(1 - \tilde{p}_{S1in}) + \tilde{p}_{S1in}(1 - c)\right] + \\ & (1 - a_S)\left[\tilde{p}_{L2out}(1 + b_S) + (1 - \tilde{p}_{L2out})(1 + b_S - c)\right] - \\ & m + i(1 - \tilde{p}_{S1in})(1 - \tilde{p}_{L2out})\end{aligned} \quad \text{(A.21)}$$

$$\begin{aligned}\tilde{w}_{S22} = {} & a_S(1 - \tilde{p}_{S1in}) + (1 - a_S)\tilde{p}_{L2out}(1 + b_S) + \\ & i(1 - \tilde{p}_{S1in})(1 - \tilde{p}_{L2out})\end{aligned} \quad \text{(A.22)}$$

Anticipated payoffs to phenotypes in group L are found by reversing all group and norm indices in the subscripts. Parameters a_S, b_S, c, m, and i are defined above.

Given that an individual either failed to coordinate or suffered a cognitive dissonance cost in the interaction phase and is therefore considering a phenotype change, a Markov process is assumed, such that the individual compares the anticipated payoff of her current phenotype against the anticipated payoff of the phenotype(s) to which she could possibly change, and biases her decision toward the phenotype with the highest mean anticipated payoff. The probability that she transitions from one phenotype to another is modeled as a logistic function (inverse logit) of the difference in anticipated payoffs between the current and potential phenotypes. A parameter $\mu \geq 0$ scales the strength of the bias for adopting the phenotype with the highest mean anticipated payoff. These probabilities for group S phenotypes are summarized as a transition matrix in Table A.3, which follows the transition rules shown in Figure A.8.

TABLE A.3. Transition probabilities between phenotypes of group S[a]

	$S11$	$S1X$	$S2X$	$S22$
$S11$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S1X})]\text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S2X})]}{P_{S11\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S11})]\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S11\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S11})]\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S11\text{tot}}}$	0
$S1X$	0	$\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]$	$\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]$	0
$S2X$	0	$\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]$	$\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]$	0
$S22$	0	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S22})]\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S22\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S22})]\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S22\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S1X})]\text{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S2X})]}{P_{S22\text{tot}}}$

[a]Probabilities of a transition from the row phenotype to the column phenotype. The total probability of a transition from $S11$ to $S11$, $S1X$, or $S22$ is the row sum $P_{S11\text{tot}} = \text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S1X})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S2X})] + \text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S11})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})] + \text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S11})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]$. $P_{S22\text{tot}}$ is the analogous sum of the numerators of the cells in the last row.

These updating assumptions are incorporated into the interaction Table A.4. Multiplying the columns Pr(self, other) by Pr($S11$) and summing over all combinations of self and other gives the frequency of phenotype $S11$ after updating. Frequencies of the other phenotypes in group S after updating are calculated analogously. Frequencies of phenotypes in group L are found by reversing all group and norm designations in Table A.4. Example recursions for $S11$ and $S1X$ are shown in Equations A.23 and A.24, giving the values of p_{S11} and p_{S1X} in the next time step.

TABLE A.4. Interaction table for individuals with phenotypes $S11$, $S1X$, $S2X$, and $S22$

Self	Other	Pr(Self, Other)[a]	$\Pr(S11)$[b]	$\Pr(S1X)$	$\Pr(S2X)$	$\Pr(S22)$
S11	S11	$a_S p_{S11}^2$	1	0	0	0
S11	S1X	$a_S p_{S11} p_{S1X}$	1	0	0	0
S11	S2X	$a_S p_{S11} p_{S2X}$	1	0	0	0
S11	S22	$a_S p_{S11} p_{S22}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S1X})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S2X})]}{P_{S11\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S11})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S11\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S11})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S11\mathrm{tot}}}$	0
S11	L11	$(1-a_S) p_{S11} p_{L11}$	1	0	0	0
S11	L1X	$(1-a_S) p_{S11} p_{L1X}$	1	0	0	0
S11	L2X	$(1-a_S) p_{S11} p_{L2X}$	1	0	0	0
S11	L22	$(1-a_S) p_{S11} p_{L22}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S1X})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S2X})]}{P_{S11\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S11})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S11\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S11})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S11\mathrm{tot}}}$	0
S1X	S11	$a_S p_{S1X} p_{S11}$	0	1	0	0
S1X	S1X	$a_S p_{S1X}^2$	0	1	0	0
S1X	S2X	$a_S p_{S1X} p_{S2X}$	0	$\frac{1}{2}+\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S1X	S22	$a_S p_{S1X} p_{S22}$	0	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S1X	L11	$(1-a_S) p_{S1X} p_{L11}$	0	1	0	0
S1X	L1X	$(1-a_S) p_{S1X} p_{L1X}$	0	1	0	0
S1X	L2X	$(1-a_S) p_{S1X} p_{L2X}$	0	$\frac{1}{2}+\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S1X	L22	$(1-a_S) p_{S1X} p_{L22}$	0	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S2X	S11	$a_S p_{S2X} p_{S11}$	0	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S2X	S1X	$a_S p_{S2X} p_{S1X}$	0	$\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\frac{1}{2}+\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S2X	S2X	$a_S p_{S2X}^2$	0	0	1	0
S2X	S22	$a_S p_{S2X} p_{S22}$	0	0	1	0
S2X	L11	$(1-a_S) p_{S2X} p_{L11}$	0	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S2X	L1X	$(1-a_S) p_{S2X} p_{L1X}$	0	$\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})\right]$	$\frac{1}{2}+\frac{1}{2}\mathrm{logit}^{-1}\left[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})\right]$	0
S2X	L2X	$(1-a_S) p_{S2X} p_{L2X}$	0	0	1	0
S2X	L22	$(1-a_S) p_{S2X} p_{L22}$	0	0	1	0
S22	S11	$a_S p_{S22} p_{S11}$	0	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S22})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S22\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S22})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S22\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S1X})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S2X})]}{P_{S22\mathrm{tot}}}$
S22	S1X	$a_S p_{S22} p_{S1X}$	0	0	0	1
S22	S2X	$a_S p_{S22} p_{S2X}$	0	0	0	1
S22	S22	$a_S p_{S22}^2$	0	0	0	1
S22	L11	$(1-a_S) p_{S22} p_{L11}$	0	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S22})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S22\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S22})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S22\mathrm{tot}}}$	$\frac{\mathrm{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S1X})]\mathrm{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S2X})]}{P_{S22\mathrm{tot}}}$
S22	L1X	$(1-a_S) p_{S22} p_{L1X}$	0	0	0	1
S22	L2X	$(1-a_S) p_{S22} p_{L2X}$	0	0	0	1
S22	L22	$(1-a_S) p_{S22} p_{L22}$	0	0	0	1

[a]Pr(interaction between Self phenotype and Other phenotype | level of in-group affinity a_S)

[b]Pr(Self=S11 after updating | current phenotype of Self, current phenotype of Other, a_S). Columns 5 through 7 interpreted analogously. $P_{S11\mathrm{tot}}$ and $P_{S22\mathrm{tot}}$ are defined in Table A.3.

$$
\begin{aligned}
p'_{S11} = a_S p_{S11} \{ & p_{S11} + p_{S1X} + p_{S2X} + \\
& p_{S22} \frac{\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S1X})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S2X})]}{P_{S11\text{tot}}} \Bigg\} + \\
& (1 - a_S) p_{S11} \{ p_{L11} + p_{L1X} + p_{L2X} + \\
& p_{L22} \frac{\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S1X})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S2X})]}{P_{S11\text{tot}}} \Bigg\}
\end{aligned}
\tag{A.23}
$$

$$
\begin{aligned}
p'_{S1X} = & a_S p_{S11} p_{S22} \frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S11})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})]}{P_{S11\text{tot}}} + \\
& (1 - a_S) p_{S11} p_{L22} \frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S11})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})]}{P_{S11\text{tot}}} + \\
& a_S p_{S1X} (p_{S11} + p_{S1X}) + \\
& a_S p_{S1X} p_{S2X} \left\{ \frac{1}{2} + \frac{1}{2} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] \right\} + \\
& a_S p_{S1X} p_{S22} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] + \\
& (1 - a_S) p_{S1X} (p_{L11} + p_{L1X}) + \\
& (1 - a_S) p_{S1X} p_{L2X} \left\{ \frac{1}{2} + \frac{1}{2} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] \right\} + \\
& (1 - a_S) p_{S1X} p_{L22} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] + \\
& a_S p_{S2X} p_{S11} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] + \\
& a_S p_{S2X} p_{S1X} \frac{1}{2} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] + \\
& (1 - a_S) p_{S2X} p_{L11} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] + \\
& (1 - a_S) p_{S2X} p_{L1X} \frac{1}{2} \text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})] + \\
& a_S p_{S22} p_{S11} \frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S22})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})]}{P_{S22\text{tot}}} + \\
& (1 - a_S) p_{S22} p_{L11} \frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S22})]\,\text{logit}^{-1}[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})]}{P_{S22\text{tot}}}
\end{aligned}
\tag{A.24}
$$

p'_{S2X} and p'_{S22} are obtained by reversing all norm subscripts in Equations A.24 and A.23, respectively.

TABLE A.5. Key for model parameters and symbols

Symbol	Range	Description
a_Z	$[0, 1]$	In-group affinity of group $Z = S$ or L: probability of attempting coordination with an in-group (as opposed to an out-group) member
m	$[0, +\infty]$	Learning cost of CCC
c	$[0, 1]$	Cognitive dissonance cost of coordinating using a non-preferred norm
b_Z	$[0, +\infty]$	Extra payoff received by a member of group $Z = S$ or L who coordinates with a member of the out-group
i	$[0, +\infty]$	Valuation of in-group identity (in currency of coordination payoffs)
μ	$[0, +\infty]$	Payoff bias in the decision to change one's phenotype
p_{Zyy}	$[0, 1]$	Frequency of a uni-cultural competence (UCC) phenotype, member of group $Z = S$ or L, who personally prefers norm y
p_{ZyX}	$[0, 1]$	Frequency of a cross-cultural competence (CCC) phenotype, member of group $Z = S$ or L, who personally prefers norm y
$\tilde{p}_{Zyin}$	$[0, 1]$	Frequency of a preference for norm y among members of group $Z = S$ or L as perceived by a fellow in-group member of group Z
$\tilde{p}_{Zyout}$	$[0, 1]$	Frequency of a preference for norm y among members of group $Z = S$ or L as perceived by a member of group not-Z
w_{Zyy}	$[0, +\infty]$	Average payoff to a UCC individual from group $Z = S$ or L, who personally prefers norm y
w_{ZyX}	$[0, +\infty]$	Average payoff to a CCC individual from group $Z = S$ or L, who personally prefers norm y
$\tilde{w}_{Zyy}$	$[0, +\infty]$	Average payoff to a UCC individual from group $Z = S$ or L, who personally prefers norm y, as perceived by a member of group Z
$\tilde{w}_{ZyX}$	$[0, +\infty]$	Average payoff to a CCC individual from group $Z = S$ or L, who personally prefers norm y, as perceived by a member of group Z

B.3. **Model Analysis.** Simulations of the trajectories and long-run equilibria for phenotype frequencies in groups S and L under a range of values for in-group affiliation (a) and the valuation of in-group identity (i) are shown in Figure 2A and B. Note that the group S-typical norm 1 can be maintained at a high frequency of both preference and use in group S in the form of the cross-culturally competent phenotype $S1X$ even when S is a minority group ($\frac{1-a_L}{1-a_S} = \frac{1}{2}$) with low bargaining power ($b_S > b_L$), whose members interact more often with the out-group than with the in-group ($a_S < \frac{1}{2}$), as long as in-group identity is sufficiently valued (reasonably large i). When i is not sufficiently large, the norm remains only in the memories of individuals ($S2X$ and $L2X$) who neither prefer nor use it (Figure 1B). Although the model represents a single generation of people who cannot forget any norm that they once knew, such an equilibrium represents the effective extinction of norm 1, as it is unlikely to be transmitted to the next generation.

Note that the phenotype frequency dynamics in group L are relatively unaffected by the equilibrium attained in group S. Group L always attains a mixed equilibrium consisting of a majority of the uni-culturally competent phenotype $L22$, and a minority of cross-culturally competent $L2X$. However, after the extinction of $S11$ in group S, there is no incentive for group L members to retain any phenotype other than $L22$, and all cross-cultural competence will likely disappear from this group in subsequent generations (not modelled). Sensitivity of phenotype dynamics to the other parameters in the model is shown in Appendix Figures A.11 and A.12.

The dynamics apparent in Figure 2A and B present an opportunity to simplify the model described above in order to make it more analytically tractable. The dynamics of interest in group S are little affected by fixing the frequency of the uni-culturally competent $L22$ phenotype in group L at 1 (compare Appendix Figures A.11 with A.14 and A.12 with A.15). Furthermore, under the parameter conditions of interest, p_{S11} is quickly lost from group S, as all $S11$ individuals transition to the cross-cultural phenotypes $S1X$ and $S2X$. Under the simplifying assumption of a constant $p_{L22} = 1$ (therefore $p_{L2X} = p_{L1X} = p_{L11} = 0$), the dynamics of p_{S11} are described by the following discrete-time difference equation, which is obtained by subtracting p_{S11} from the recursion in Equation A.23.

$$\Delta p_{S11} = \frac{p_{S11}[-1 + a_S(p_{S11} + p_{S1X} + p_{S2X})](1 + e^v + e^w + e^x)}{1 + e^v + e^w + e^x + e^y + e^z} \tag{A.25}$$

where the exponents $\{v, w, x, y, z\} \in \mathbb{R}$ are themselves each functions of model parameters. Because $a_S(p_{S11} + p_{S1X} + p_{S2X})$ is always < 1, p_{S11} inevitably decreases to 0 regardless of the values of the exponents. This is a consequence of the modelling assumption that prohibits all transitions from cross-culturally competent to uni-culturally competent phenotypes (i.e., individuals cannot unlearn a norm). As we are particularly interested in the frequencies of the cross-cultural phenotypes at equilibrium, we can simplify the model further by examining the system dynamics after p_{S11} reaches 0.

Under the simplifying assumption that $p_{L2X} = p_{L1X} = p_{L11} = p_{S11} = 0$, the system dynamics can be represented by a single difference equation, found by subtracting p_{S1X} from Equation A.24. (For reference, Table A.5 provides a key for symbols in the following equations.)

$$\Delta p_{S1X} = a_S p_{S1X} p_{S2X} \frac{1}{2} P - p_{S1X}\left[a_S p_{S2X}\frac{1}{2} + a_S(1 - p_{S1X} - p_{S2X}) + (1 - a_S)\right](1 - P) \tag{A.26}$$

where

$$P = \frac{1}{1 + e^{-\mu[iF + c(2a_S F - 1)]}} \tag{A.27}$$

is the probability of a transition $S2X \to S1X$ conditional on $S2X$ and $S1X$ coordinating on norm 1, and where

$$F = p_{S1X}\left(p_{S1X} + \frac{1}{2}p_{S2X}\right) + p_{S2X}\left(\frac{1}{2}p_{S1X}\right) = p_{S1X}(p_{S1X} + p_{S2X}) \tag{A.28}$$

is the probability with which norm 1 is observed to be used in group S, and is thus the effective frequency of norm 1 in group S as perceived by group S members.

The first term in Equation A.26 is the probability of an interaction between $S1X$ and $S2X$ individuals, multiplied by the probability that they coordinate using norm 1 (i.e, $\frac{1}{2}$), multiplied by the probability that $S2X$ transitions to $S1X$ conditional on such an interaction (i.e.,

P). During such an interaction, only $S2X$ would consider changing her preferred norm, as only she suffers the cognitive dissonance cost. Thus, conditional on such an interaction, the probability of a transition $S1X \to S2X$ is 0 rather than $1 - P$. The second term is the probability that $S1X$ coordinates with any in-group or out-group member using norm 2, multiplied by the probability that $S1X$ transitions to $S2X$ conditional on such an interaction (i.e., $1 - P$). Analogous to the previous condition, during such an interaction, only $S1X$ would consider changing her preferred norm, as only she suffers the cognitive dissonance cost. Note that $p_{S22} = 1 - p_{S1X} - p_{S2X}$ and all members of the out-group, with whom interaction occurs with probability $1 - a_S$, are phenotype $L22$.

P is a logistic function with range $(0, 1)$, increasing over its domain. The inflection point is located at $x = -c(2a_S F - 1)$, at which the slope is μ, the strength of the bias for adopting the phenotype with the highest mean anticipated payoff. For all $x >$ the inflection point, the probability of a transition $S2X \to S1X$ is $> \frac{1}{2}$. Note that the inflection point moves left with increasing in-group affinity (a_S) and increasing probability of observing the use of norm 1 within the in-group (F). The cost of cognitive dissonance for coordination using a non-preferred norm (c) scales this effect. For instance, all else being equal, when many interactions occur with members of the out-group L (all of whom use norm 2) and/or norm 1 is used infrequently within the in-group S ($a_S F < \frac{1}{2}$), increasing c decreases the probability of $S2X \to S1X$ (i.e., by moving the inflection point right). However, when norm 1 is used frequently within the in-group, and many interactions occur there ($a_S F > \frac{1}{2}$), increasing c increases this probability (by moving the inflection point left). This effect is due to the fact that cross-cultural individuals who prefer a norm that is rarely used during interactions suffer a higher average cost of cognitive dissonance, as they must interact more frequently using their non-preferred norm. A central assumption of the model is that individuals modify their phenotype (or not) in order to minimize such costs in the future. An increase in c therefore increases the probability that an individual who prefers a norm that is rare among her interaction partners will transition to a phenotype that prefers a more common norm in order to reduce this cost.

The argument of the logistic function is $x = iF$. When this argument is greater than the domain value of the inflection point, the probability of $S2X \to S1X$ is $> \frac{1}{2}$. Note that the value of the argument increases with increasing valuation of cultural identity (i) and increasing probability of observing the use of norm 1 within the in-group (F). This effect reflects the assumption that an individual who values her in-group cultural identity receives utility from preferring a particular cultural norm only to the extent that this norm is perceived to be both common within the in-group and rare within the out-group (recall that in this simplified model, norm 1 is absent in the out-group). If an individual places a high value on cultural identity, she will be more likely to adopt a phenotype that prefers a norm common in the in-group when this norm is absent (or rare) in the out-group.

Note that, given the simplifying assumptions above, $\Delta p_{S22} = 0$, as there is no incentive for $S22$ to transition to a cross-culturally competent phenotype when there are no longer any uni-culturally competent individuals who prefer norm 1 (i.e., in the simplified model, $p_{S11} = 0$). Similarly, $\Delta p_{S2X} = -\Delta p_{S1X}$, as the dynamics now comprise only transitions between the two cross-culturally competent states $S1X$ and $S2X$. Equation A.26 shows that the dynamics of the simplified system are affected neither by the power difference between groups S and L (i.e., b_S and b_L), nor by the cost of learning and maintaining cross-cultural competence (m). This is because both $S1X$ and $S2X$ always successfully coordinate with the out-group and receive the inter-group interaction payoff bonus b_S, and both suffer the cross-cultural learning cost m. Therefore, individuals' phenotype transition decisions, which are based on the difference in average anticipated payoffs between these phenotypes, cannot depend on b_S and m.

B.3.1. *Equilibria.* Setting Equation A.26 equal to 0, it can be seen that this system has equilibria at $p_{S1X} = 1$ when $a_S = 1$, and $p_{S1X} = 0$ for all a_S. For the remaining analysis, I focus on the condition when $a_S < \frac{1}{2}$. This corresponds to a context in which members of a disempowered minority group interact more frequently with members of a powerful majority group than they do with their fellow in-group members, and this is the context in which it is most difficult to sustain the minority-typical norm. Sustaining the minority group S-typical norm 1 occurs only when the cross-cultural phenotype $S1X$ is retained in group S at equilibrium. An equilibrium that includes only phenotypes $S2X$ and $S22$ corresponds to a situation in which norm 1 is preferred by no one, is never used, and is unlikely to be passed on to future generations.

Equation A.26 reveals that, when individuals interact more frequently with the out-group than with the in-group ($a_S < \frac{1}{2}$) and do not value cultural group identity ($i = 0$), phenotype $S1X$ is always lost from the population at equilibrium. In this case, the argument of the logistic function in Equation A.27 is 0, and the inflection point is always > 0, constraining $P < \frac{1}{2}$. Thus, the second term in Equation A.26 is always greater in absolute value than the first, and Δp_{S1X} is negative until p_{S1X} reaches 0.

However, if individuals place sufficient value on their cultural group identity ($i > 0$), mixed equilibria containing $S1X$ are possible. I have been unable to find a closed-form equilibrium solution to the above dynamics by solving for the state variables p_{A1X} and p_{A2X}. However, setting Equation A.26 equal to zero and solving for i yields equilibrium values of p_{A1X} and p_{A2X} whenever the following holds:

$$i = \frac{\frac{1}{\mu}\ln\left[\frac{2(1-a_S p_{S1X})}{a_S p_{S2X}} - 1\right] - c(2a_S F - 1)}{F} \tag{A.29}$$

where

$$F = \ p_{S1X}(p_{S1X} + p_{S2X}) \tag{A.30}$$

is the probability with which norm 1 is observed to be used within group S. Note that the degree to which an individual values the cultural identity of her in-group, i, is defined to be ≥ 0. From Equation A.29 it can be seen that all equilibria including $S1X$ are mixed equilibria: for any finite i, $p_{S1X} < 1$ because $p_{S2X} > 0$ (from denominator of the first term in the logarithm).

Equation A.29 reveals that, for a given i of sufficient size, there may be two equilibrium values of p_{S1X}, as there are two ways to increase the right side of the equation to match any i. One of these equilibria is reached as p_{S1X} approaches 1 (the denominator of the first term in the logarithm approaches 0). The second is reached as p_{S1X} approaches 0, making the denominator of Equation A.29 approach 0. It can be shown numerically that, when two distinct mixed equilibria exist, one of them is unstable, defining the basin of attraction for the other stable equilibrium. Figure 2C shows this basin of attraction for three values of i, and a range of values of p_{S22}. Note that, in the absence of $S11$ in this simplified model, p_{S22} is not a state variable, as its value is constant across time steps. For a given proportion of $S22$ individuals in group S, the frequencies of $S1X$ and $S2X$ that fall within the basin of attraction of a stable mixed equilibrium compose the line segment (parallel to the trajectory arrows) connecting a point on the blue line to the $S1X - S22$ axis. The frequencies of $S1X$ and $S2X$ at this stable equilibrium correspond to the point where the segment crosses the red line. As i increases, so also increase the maximum values of p_{S2X} and

p_{S22} that can still result in a stable equilibrium which includes $S1X$ (the blue and black points, respectively, in Figure 2C).

Equation A.29 also reveals the fundamental antagonism between group identity valuation (i) and cognitive dissonance (c). When interaction with the out-group occurs more frequently than interaction with the in-group ($a_S < \frac{1}{2}$), the product $2a_S F$ is less than 1, making the second term in the numerator positive. Thus, all else being equal, a larger value of c requires a larger value of i in order to achieve a mixed equilibrium containing $S1X$. In other words, to maintain the minority-typical norm at equilibirum in the form of $S1X$ when intergroup interaction is frequent, valuation of group identity must outweigh the cost of cognitive dissonance from having to coordinate often using one's non-preferred norm.

B.3.2. *Resistance to Uni-Cultural Invasion.* In this model, both the full and simplified versions, an important assumption is that individuals cannot unlearn or forget a norm that they know. This prevents all transitions from cross-cultural phenotypes to uni-cultural phenotypes (see Figure A.8). However, it may be the case that parents, when deciding whether to teach certain norms to their children, weigh the anticipated costs and benefits (i.e., average anticipated payoffs), both to themselves and to their children (Bisin and Verdier 2001). In particular, if cross-culturally competent parents perceive that the anticipated payoff to a uni-cultural phenotype outweighs the payoff to a cross-cultural phenotype, they may make the strategic decision to transmit to their children only one of the two norms that they know. Similarly, children themselves may be less receptive to learning two norm variants (as opposed to one), if they perceive that the cost of learning the second norm outweighs the benefit that knowing this norm would afford them. Both of these processes may contribute to the well-documented inter-generational loss of minority languages in situations of language contact (Portes and Rumbaut 2014) and distinctive norms in the context of immigration (Gans 1979). Thus, in theory, a norm typical of a disempowered minority group may attain high frequency through cross-cultural competence in a mixed equilibrium within any one generation, yet disappear in the next, or subsequent, generations. An informative representation of the inter-generational dynamics of such phenotypes requires an age-structured model, incorporating assumptions about demographic processes, marriage assortment, child socialization, and teaching/learning strategies. Such a model is beyond the scope of the present study. The goal of the following analysis is simply to find the conditions under which cross-cultural $S1X$ individuals, acting only according to what they perceive as their anticipated children's best interests, would prefer to create $S1X$ rather than $S22$ or $S2X$ offspring, given conditions in the next generation identical to those in the present.

Within the basin of attraction for stable mixed equilibria containing $S1X$, shown in Figure 2C, the average perceived payoff to $S1X$ exceeds that to $S2X$, i.e., $\tilde{w}_{S1X} > \tilde{w}_{S2X}$. Thus, within this region, $S1X$ parents are expected to create $S1X$ offspring, rather than $S2X$ or $S22$, as long as $\tilde{w}_{S1X}$ is greater than $\tilde{w}_{S22}$. Solving the inequality $\tilde{w}_{S1X} > \tilde{w}_{S22}$ yields the threshold value of i, above which the $S1X$ phenotype is expected to be favored over $S22$ in the next generation, all else being equal:

$$i > \frac{m + c(1 - a_S F) - a_S F}{F} \tag{A.31}$$

where

$$F = \; p_{S1X}(p_{S1X} + p_{S2X}) \tag{A.32}$$

as in Equation A.29.

As can be seen in Equation A.31, parents' decision to produce $S1X$ offspring is expected to be sensitive to the learning cost (m) that such cross-cultural children are likely to incur. When m is high, the value placed on cultural identity (i) must also be high in order to prevent $S1X$ parents from teaching their children only a single norm, and thereby raising uni-culturally competent offspring. Similarly, when parents anticipate that their children will have a high probability of interacting with others who prefer norm 2, i.e., $1 - a_S F$, and the cost of cognitive dissonance (c) is high, they will likely teach their children to prefer norm 2 in order to insulate them from this cost, unless i is sufficiently high. Thus, when m and c are high, parents should prefer to raise $S22$ children. Conversely, the higher the frequency of $S1X$ at equilibrium and/or the more interaction occurs with in-group members (high a_S), the lower must i be in order for parents to favor $S1X$ in the next generation. Note that in this simplified model, the power difference between groups (b_S and b_L) has no effect on parents' decisions, as all phenotypes that they would consider for their offspring ($S1X$, $S2X$, and $S22$) can successfully coordinate with all possible interaction partners.

Figure 2C shows, for given values of i and m (and the other model parameters), the range of phenotype frequencies such that $\tilde{w}_{S1X} > \tilde{w}_{S22}$ (grey regions). When phenotype frequencies fall within the intersection of this grey region and the basin of attraction for the mixed equilibria, parents would be expected to prefer $S1X$ offspring. Note that, when the learning cost of cross-cultural competence (m) is high, not all stable equilibrium frequencies of $S1X$ (red line) are expected to result in the inter-generational transmission of the $S1X$ phenotype. Thus, while the sustainability of a minority-typical norm within one generation (in the form of $S1X$) does not depend on m, the sustainability of this norm across generations may.

Figures A.9 and A.10 show the separate (and opposing) effects of i and m on the range of phenotype frequencies (grey regions) corresponding to $\tilde{w}_{S1X} > \tilde{w}_{S22}$. As apparent in Inequality A.31, increasing identity valuation i increases the range of situations (group phenotype frequencies) in which parents would prefer to create $S1X$, rather than $S22$ (or $S2X$), children. However, increasing the cost m of learning a new norm decreases the range of situations in which parents would prefer to create $S1X$, rather than $S22$, children.

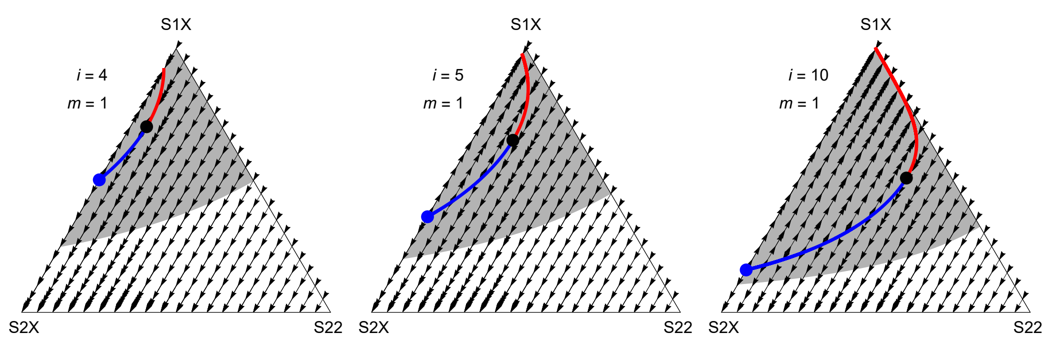

FIGURE A.9. Ternary plots for a simplified model ($p_{S11} = 0$, $p_{L22} = 1$), with parameter conditions, and interpretation, analogous to Figure 2C, showing the effect of increasing values of identity valuation (i) on the phenotype frequencies (area in grey) in which anticipated payoffs $\tilde{w}_{S1X} > \tilde{w}_{S22}$, while holding the learning cost m constant. Increasing i increases the range of phenotype frequencies for which Inequality A.31 holds. The intersection of the grey region and the basin of attraction (region bound by the blue line and the $S1X - S22$ axis) is the set of phenotype frequencies for which inter-generational transmission of $S1X$ is plausible.

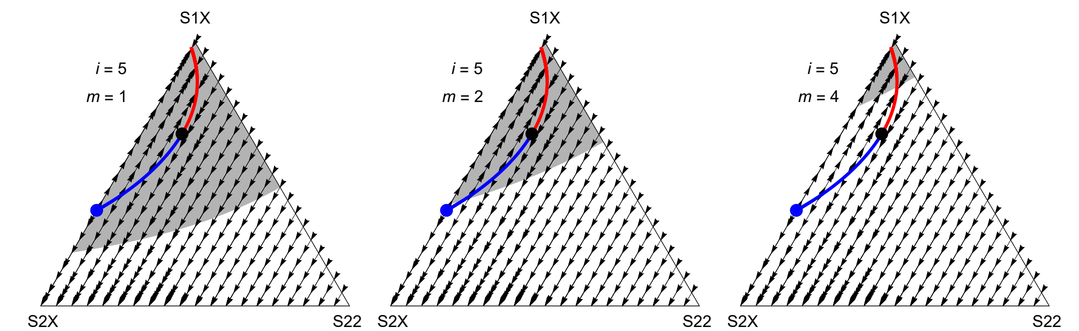

FIGURE A.10. Ternary plots analogous to Figure A.9, showing the effect of increasing values of the norm-learning cost m on the phenotype frequencies (area in grey) in which anticipated payoffs $\tilde{w}_{S1X} > \tilde{w}_{S22}$, while holding identity valuation i constant. Given a constant i, increasing m decreases the range of phenotype frequencies for which Inequality A.31 holds.

B.3.3. *Model Simplification and Sensitivity.* Figures A.11 and A.12 show phenotype trajectories and long-run phenotype frequencies, respectively, derived from simulations of the full theoretical model under a range of parameter conditions. Figures A.14 and A.15 show the same parameter conditions for the simplified model where phenotypes in the majority group L do not evolve, and p_{L22} is fixed at 1. The similarity of the two sets of figures suggests that many aspects of the dynamics of the full model are accurately approximated by the simplified model, justifying the analysis strategy in the main text. Figures A.16 and A.17 show the same parameter conditions for a situation in which group L is five times larger than group S ($\frac{1-a_L}{1-a_S} = \frac{1}{5}$). In all other simulations in this manuscript, group L is twice as large as group S. Comparing these plots to those above, such a change in relative group size tends to speed up the dynamics, but does not qualitatively affect the sensitivity of the model to changes in other parameters. Thus, the conclusions reported in the main text appear robust to reasonable changes in the relative sizes of the groups.

We can explore the effect of different parameters on the likelihood of sustaining a preference for S-typical norm 1 at equilibrium in minority group S, either in the form of the uni-culturally competent phenotype $S11$ (black) or the cross-culturally competent phenotype $S1X$ (red). It is readily observed that, as long as there is some inter-group interaction (in-group affinity $a < 1$), $S11$ is invariably lost from the population, though this loss is slower when a is large (A.11A and B) or the learning cost of cross-cultural competence, m is large (A.11G and A.12E). Thus, as a result of the assumptions of the model, minority norm 1 can be sustained only in the form of the cross-culturally competent phenotype $S1X$. When group L is greater in size than group S (operationalized as $a_L > a_S$), there is more than a minimal amount of inter-group interaction (a_S is not extremely close to 1), and no particular value is placed on group identity ($i = 0$), $S1X$ is always lost from the population, even in the absence of power differences between groups (A.11A and A.12A). In order for $S1X$ to be maintained at equilibrium, the maximum frequency attained by $S1X$ must be sufficiently high to fall within the basin of attraction of a stable mixed equilibrium (Figure 2C). Thus, parameters that, when increased, increase the maximum height of the $S1X$ trajectory (red) in Figure A.11, or the area of long-run $S1X$ dominance (red) in Figure A.12, contribute positively to the sustainability of minority norm 1.

Figures A.11B and A.12B show that, as long as group L is larger than group S, the sustainability of $S1X$ is only marginally affected by the fact that group L has higher power than group S ($b_S > b_L$). Furthermore, it is more difficult to sustain $S1X$ when inter-group interaction is high (low a), and easier when group identity is valued (Figures A.11C and D).

For given values of a and i, even very large differences in bargaining power ($b_S >> b_L$) have little effect on the sustainability of $S1X$ (Figures A.11E and A.12C). As explained in the main text, this is because the most important phenotype dynamics in group S involve the two cross-culturally competent phenotypes $S1X$ and $S2X$, both of which are equally affected by group-level differences in bargaining power. Increasing b_S increases the rate of decrease of $S11$ by increasing the perceived average payoff difference between $S11$ (which does not receive b_S when interacting with $L22$) and $S1X$ and $S2X$ (both of which always receive b_S during inter-group interaction). However, changing b_S does not affect the perceived average payoff difference between $S1X$ and $S2X$.

The cognitive dissonance cost associated with coordinating using a non-preferred norm (c) is suffered most by the phenotype whose preferred norm is preferred by a minority of the pool of her potential interaction partners (including members of both the in- and out-group). As long as norm 1 is preferred by the majority of S members, while norm 2 is preferred by the majority of L members, increasing inter-group interaction (lowering a_S) increases the burden of c for $S1X$ relative to

$S2X$ individuals. Increasing a_S has the opposite effect. This can be seen in Figures A.11F and A.12D.

As explained in the main text, increasing the learning cost of cross-cultural competence (m) has only an indirect effect on the dynamics of $S1X$ and $S2X$, both of which suffer this cost equally. The effect of a large m is to decrease the rate at which the uni-culturally competent phenotype $S11$ transitions to one of these two cross-culturally competent phenotypes (Figure A.11G). Figures A.11G and A.12E, show that, at low values of i, large m is detrimental to the sustainability of $S1X$. This is because, while large m decreases the rate at which $S11$ transitions to $S1X$ (which suffers m), it does not decrease the rate at which $S1X$ transitions to $S2X$ (both of which suffer m). When i is sufficiently large and norm 1 is perceived to be be at high frequency in group S, transitions from $S1X$ to $S2X$ are slowed enough to counteract this effect.

Increasing the degree of bias toward higher perceived payoffs in individuals' decisions to modify their phenotype (μ) increases the sustainability of $S1X$ (Figures A.11H and A.12F). This is because, at early stages of the dynamics when $S11$ is still common, $S1X$ has a higher perceived average payoff than $S2X$ because norm 1 is perceived to be in the majority. High μ decreases the probability (assumed to be non-zero) that $S11$ or $S1X$ transitions to $S2X$ despite the fact that $S2X$ is perceived to have a lower average payoff than either of these phenotypes.

Note that all of the parameter modifications examined above have only negligible effects on the dynamics of majority group L, which invariably attains a mixed equilibrium of predominantly $L22$ and $L2X$, and possibly a low frequency of $L1X$ (e.g., Figures 2A and 1A). As explained in the main text, after $S11$ is lost from group S, there is no incentive for cross-culturally competent group L parents (i.e, $L2X$ and $L1X$) to teach anything other than norm 2 to their children. This is because all group S individuals with whom their children might interact are cross-culturally competent, and will coordinate using norm 2 with a uni-culturally competent individual who does not know norm 1. Thus, in this context, after one generation, group L will consist entirely of $L22$ individuals.

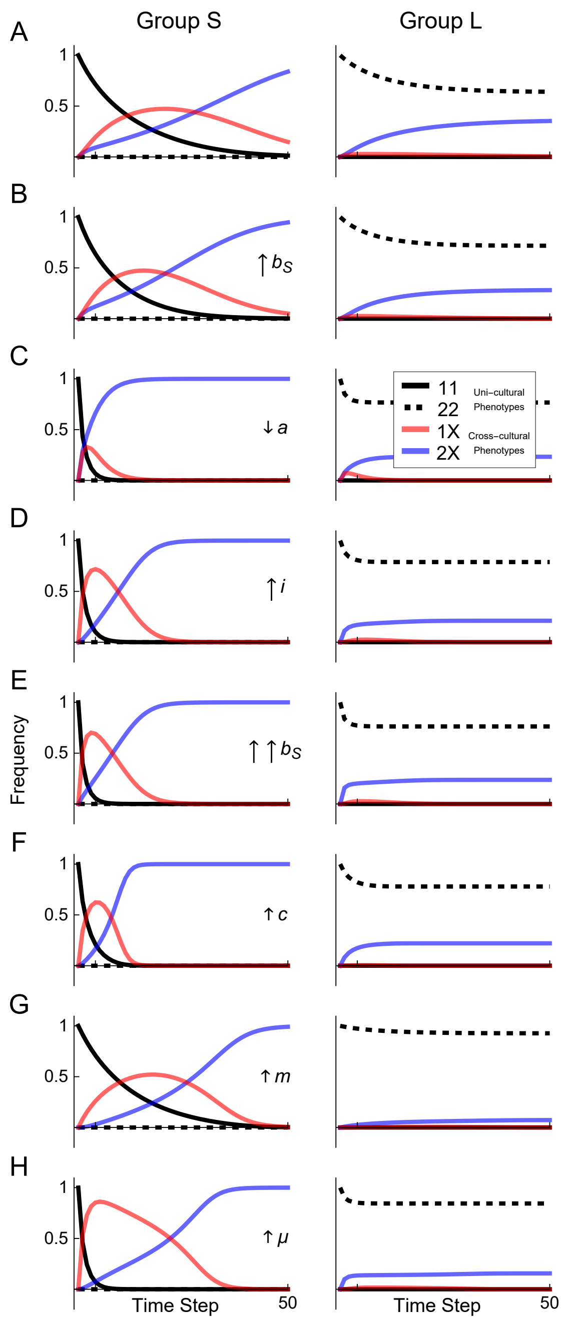

FIGURE A.11. (Caption next page.)

FIGURE A.11. Sensitivity of phenotype trajectories for the minority group (S) and majority group (L), simulating from the full model. **A)** No power difference between groups, high in-group affinity: $(b_S, b_L) = (0, 0)$, $(a_S, a_L) = (0.8, 0.9)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, $i = 0$, and initial phenotype frequencies $p_{S11} = p_{L22} = 1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Low affinity: sames as **B**, except $(a_S, a_L) = (0.4, 0.7)$. **D)** Identity valuation: sames as **C**, except $i = 1$. **E)** Large power difference: sames as **D**, except $(b_S, b_L) = (5, 1)$. **F)** High cognitive dissonance: sames as **D**, except $c = 0.9$. **G)** High cross-cultural learning cost: same as **D**, except $m = 2$. **H)** High payoff-bias for copying: sames as **D**, except $\mu = 3$.

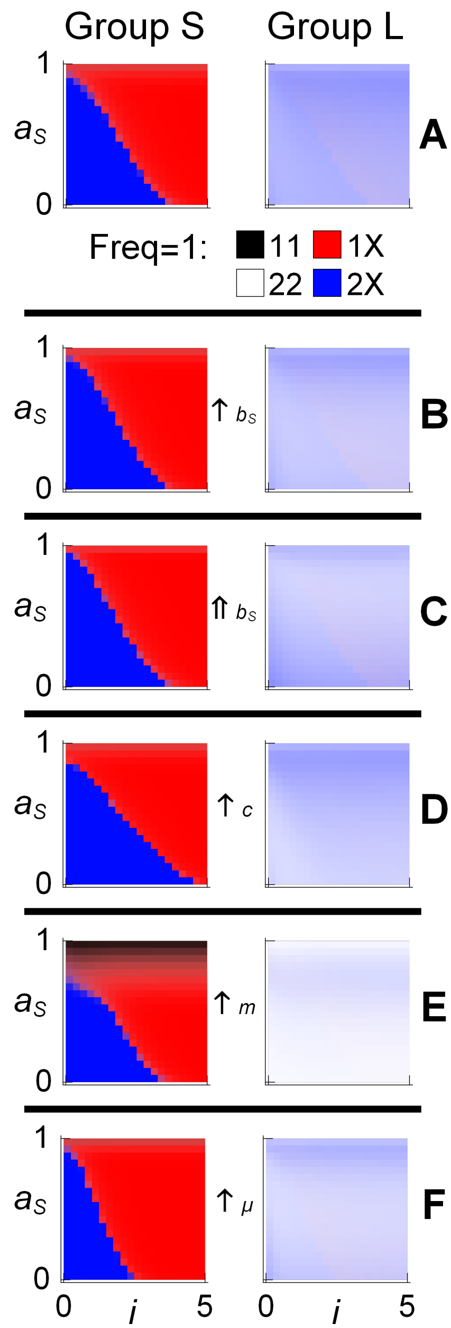

Figure A.12. (Caption next page.)

FIGURE A.12. (Figure on previous page) Sensitivity of long-run phenotype frequencies for the minority group (S) and majority group (L), simulating from the full model for 100 time steps. **A)** No power difference between groups: $(b_S, b_L) = (0, 0)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, and initial phenotype frequencies $p_{S11} = p_{L22} = 1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Large power difference: sames as **B**, except $(b_S, b_L) = (5, 1)$. **D)** High cognitive dissonance: sames as **B**, except $c = 0.9$. **E)** High cross-cultural learning cost: sames as **B**, except $m = 2$. **F)** High payoff-bias for copying: same as **B**, except $\mu = 3$.

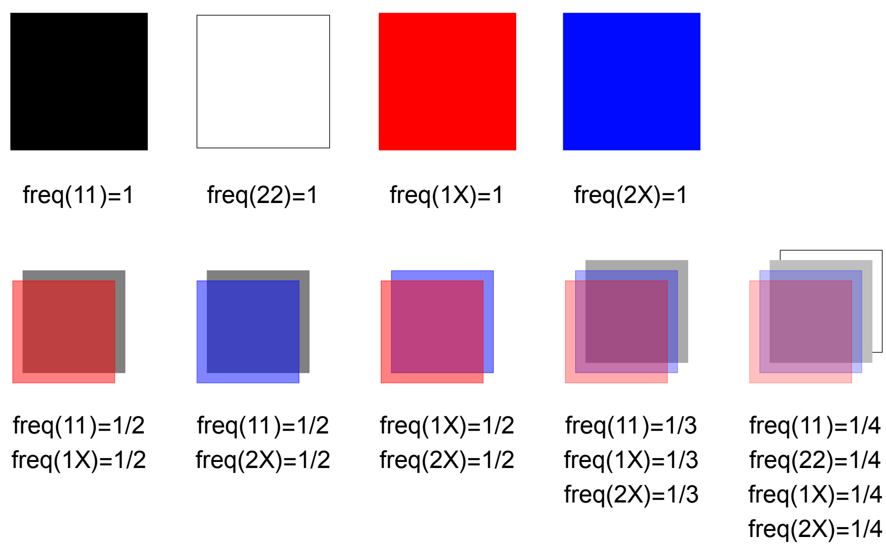

FIGURE A.13. Examples of color mixing as representations of phenotype frequencies in density plots such as Figures 2B (main text) and A.12 (above).

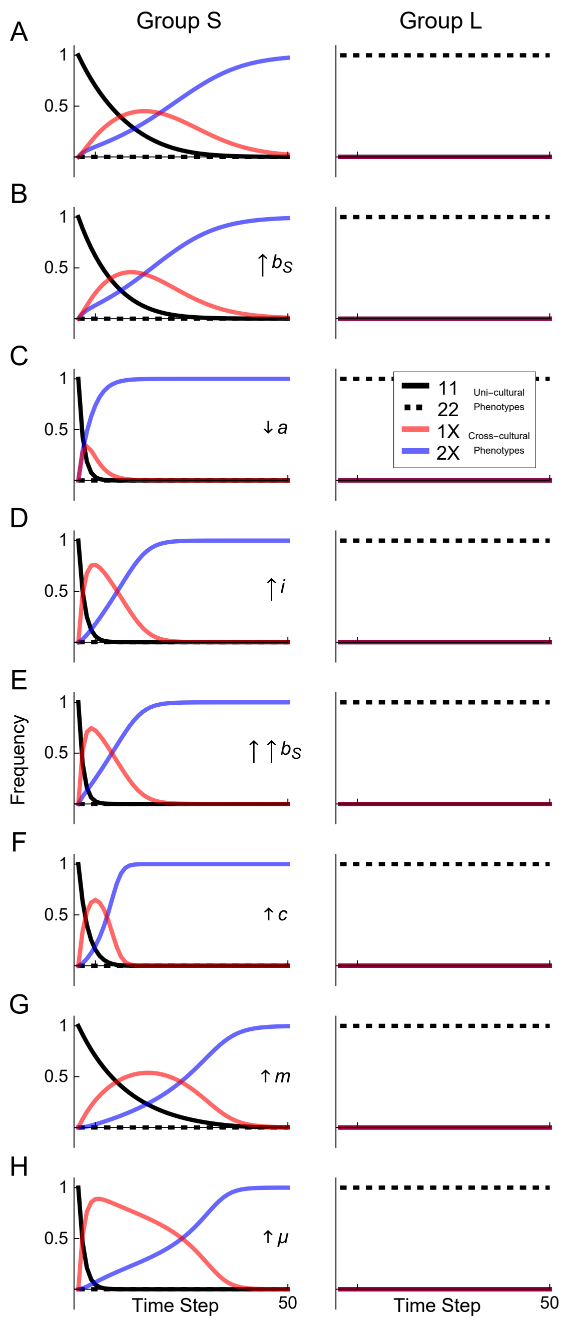

FIGURE A.14. (Caption next page.)

FIGURE A.14. Sensitivity of phenotype trajectories for the minority group (S) and majority group (L), simulating from the simplified model where p_{L22} is permanently fixed at 1. **A)** No power difference between groups, high in-group affinity: $(b_S, b_L) = (0, 0)$, $(a_S, a_L) = (0.8, 0.9)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, $i = 0$, and initial phenotype frequencies $p_{S11} = p_{L22} = 1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Low affinity: sames as **B**, except $(a_S, a_L) = (0.4, 0.7)$. **D)** Identity valuation: sames as **C**, except $i = 1$. **E)** Large power difference: sames as **D**, except $(b_S, b_L) = (5, 1)$. **F)** High cognitive dissonance: sames as **D**, except $c = 0.9$. **G)** High cross-cultural learning cost: sames as **D**, except $m = 2$. **H)** High payoff-bias for copying: same as **D**, except $\mu = 3$.

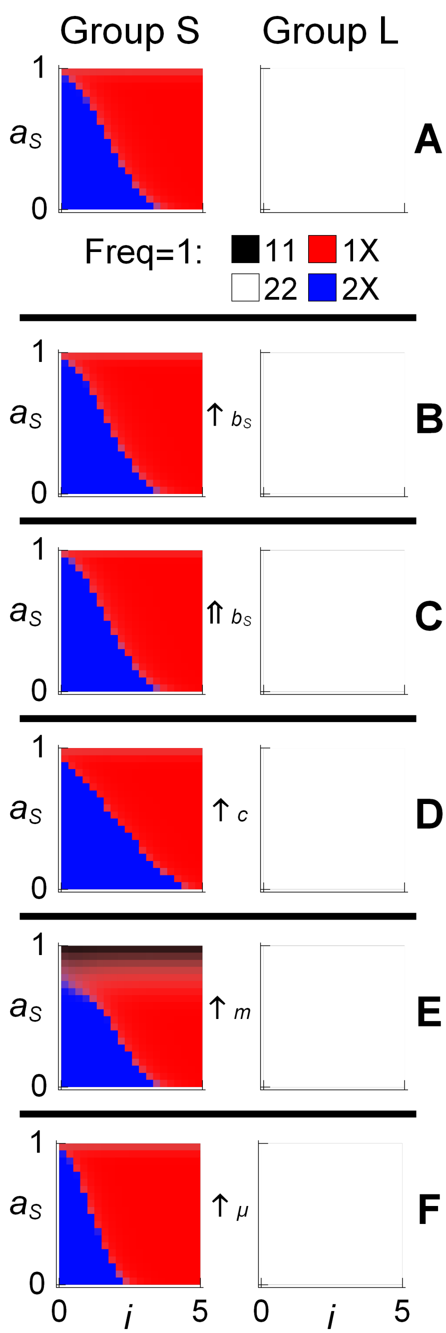

FIGURE A.15. (Caption next page.)

FIGURE A.15. Sensitivity of long-run phenotype frequencies for the minority group (S) and majority group (L), simulating from the simplified model (p_{L22} fixed at 1) for 100 time steps. **A)** No power difference between groups: $(b_S, b_L) = (0, 0)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, and initial phenotype frequencies $p_{S11} = p_{L22} = 1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Large power difference: sames as **B**, except $(b_S, b_L) = (5, 1)$. **D)** High cognitive dissonance: sames as **B**, except $c = 0.9$. **E)** High cross-cultural learning cost: sames as **B**, except $m = 2$. **F)** High payoff-bias for copying: same as **B**, except $\mu = 3$.

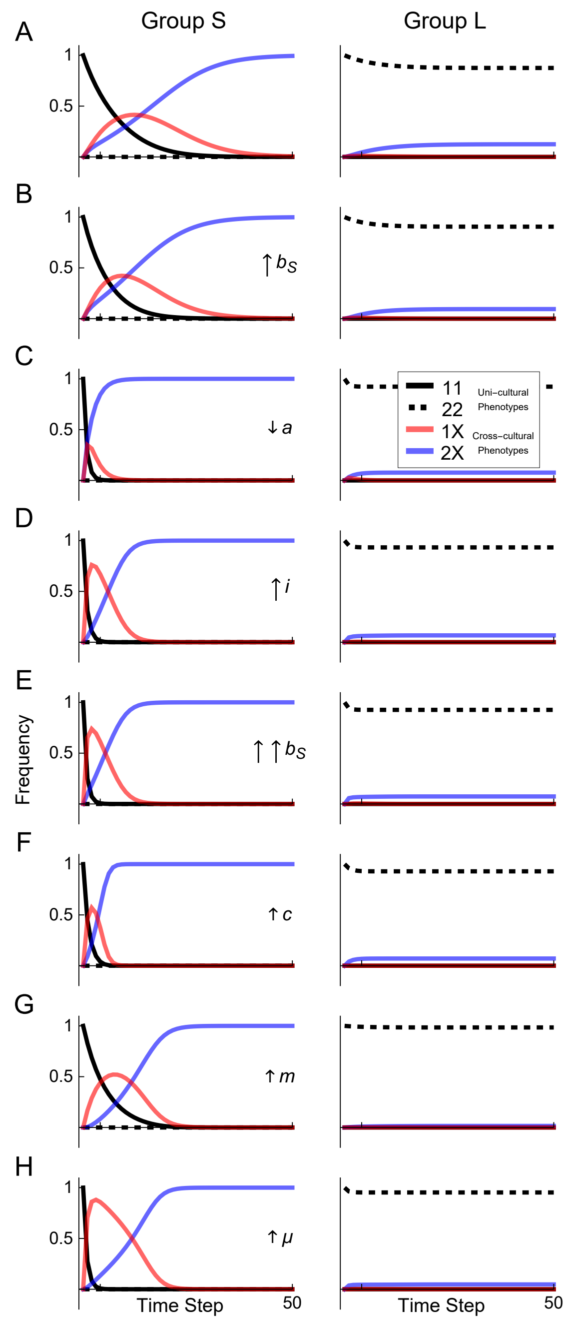

FIGURE A.16. (Caption next page.)

FIGURE A.16. Sensitivity of phenotype trajectories for the minority group (S) and majority group (L), simulating from the full model, where group L is five times larger than group S. In all other simulations, group L is twice as large as group S. **A)** No power difference between groups, high in-group affinity: $(b_S, b_L) = (0, 0)$, $(a_S, a_L) = (0.75, 0.95)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, $i = 0$, and initial phenotype frequencies $p_{S11} = p_{L22} = 1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Low affinity: sames as **B**, except $(a_S, a_L) = (0.25, 0.85)$. **D)** Identity valuation: sames as **C**, except $i = 1$. **E)** Large power difference: sames as **D**, except $(b_S, b_L) = (5, 1)$. **F)** High cognitive dissonance: sames as **D**, except $c = 0.9$. **G)** High cross-cultural learning cost: same as **D**, except $m = 2$. **H)** High payoff-bias for copying: sames as **D**, except $\mu = 3$.

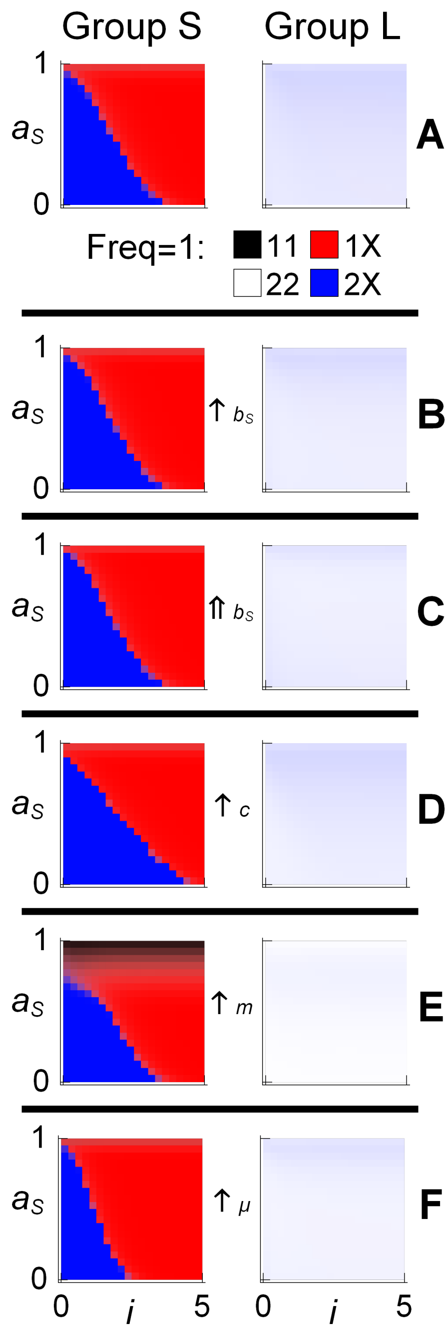

FIGURE A.17. (Caption next page.)

FIGURE A.17. Sensitivity of long-run phenotype frequencies for the minority group (S) and majority group (L), simulating from the full model for 100 time steps, where group L is five times larger than group S. In all other simulations, group L is twice as large as group S. **A)** No power difference between groups: $(b_S, b_L) = (0, 0)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, and initial phenotype frequencies $p_{S11} = p_{L22} = 1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Large power difference: sames as **B**, except $(b_S, b_L) = (5, 1)$. **D)** High cognitive dissonance: sames as **B**, except $c = 0.9$. **E)** High cross-cultural learning cost: sames as **B**, except $m = 2$. **F)** High payoff-bias for copying: same as **B**, except $\mu = 3$.

B.4. **Relating the Theoretical Model to Empirical Data.**

B.4.1. *Estimation of Model Parameters.* Comparing Figure 1A and B, it can be seen that the empirically-estimated frequencies of the four norm phenotypes among Matsigenka and Mestizos, at the time data were collected, approximately correspond to a point early in the modeled pre-equilibrium dynamics (before time step 5), given a set of assumptions about the values of the theoretical model parameters a (in-group affinity), m (learning cost of CCC), c (cognitive dissonance cost), b (additional out-group coordination payoff), i (valuation of in-group identity), and μ (payoff bias in norm adoption), as well as initial phenotype frequencies. Importantly, it can be seen that the estimated phenotype frequencies among Matsigenka who have more inter-ethnic experience (education among Mestizos) differ substantially from estimated phenotype frequencies among all Matsigenka (top versus bottom row of Figure 1A). These differences correspond to the same point in the modeled dynamics, given a reduction in the value of the parameter a (corresponding to a higher probability of inter-group interaction) (bottom row of Figure 1B). The fact that changing the parameter (a) in the theoretical model controlling inter-group interaction results in model predictions that match the change in empirically-estimated phenotype frequencies among individuals who engage in more inter-group interaction, serves to increase confidence that the model represents processes at work in the real world.

However, ideally, we would like a more objective way of determining the combination of parameter values in the theoretical model that lead to predictions that best match the empirical phenotype frequencies. Here I present a preliminary strategy using Bayesian estimation of model parameters conditional on the data (observed phenotype frequencies). Given arbitrarily-chosen initial phenotype frequencies and an arbitrarily-chosen time-step in the dynamics of the full model (expressed as the recursions in Equations A.23 and A.24), I use a Hamiltonian Monte Carlo sampling engine (Stan Development Team 2018) to estimate posterior probability distributions for the parameters a, m, c, b, i, and μ, conditional on observed phenotype frequencies among all Matsigenka and Mestizos, and among those with particular inter-group interaction experiences.

Interviewees' phenotypes are assigned based on their personally-preferred norm and their in- and out-group guesses (Table A.1), under the simplifying assumption that most Matsigenka personally prefer norm 1 and most Mestizos personally prefer norm 2. Any interviewee with a phenotype other than $1X$, $2X$, 11, and 22 is removed, as the theoretical model makes predictions only about these four phenotypes. The phenotype of interviewee j belonging to ethnic group $x = S$ or L, $y_{x,j}$, is modeled using a Categorical (i.e., Multinomial) likelihood, with the probabilities of each of the four categories equated with the predicted frequencies of the four phenotypes after a given number of recursions ($tmax$) of the theoretical model. Reasonably uninformative priors are placed on the theoretical model's parameters:

$$y_{S,j} \sim \text{Categorical}(p_{S11_{tmax}}, p_{S1X_{tmax}}, p_{S2X_{tmax}}, p_{S22_{tmax}}) \tag{A.33}$$
$$y_{L,j} \sim \text{Categorical}(p_{L11_{tmax}}, p_{L1X_{tmax}}, p_{L2X_{tmax}}, p_{L22_{tmax}}) \tag{A.34}$$
$$a_L \sim \text{Uniform}(0, 1) \tag{A.35}$$
$$a_S = 1 - 2(1 - a_L) \tag{A.36}$$
$$m \sim \text{Exponential}(1) \tag{A.37}$$
$$c \sim \text{Uniform}(0, 1) \tag{A.38}$$
$$b_L \sim \text{Exponential}(1) \tag{A.39}$$
$$b_{more} \sim \text{Exponential}(1) \tag{A.40}$$
$$b_S = b_L + b_{more} \tag{A.41}$$
$$i \sim \text{Exponential}(1) \tag{A.42}$$
$$\mu \sim \text{Exponential}(1) \tag{A.43}$$

The parameter a_S is defined such that Group L is twice the size of Group S, and b_S is defined such that Group S has lower bargaining power (higher b, see Appendix B.2.1). As in the main text, Group S comprises Matsigenka and Group L comprises Mestizos.

Parameter estimation was accomplished with RStan 2.17.3 (Stan Development Team 2018), running four Hamiltonian Monte Carlo chains in parallel until convergence was suggested by a high effective number of samples (> 500) and $\hat{R}$ estimates of 1.00 (McElreath 2016, pg 257). This entailed 2000 samples per chain, half of which were warm-up. Data and analysis scripts in R (R Core Team 2017) implementing RStan are available from Github at `https://github.com/jabunce/Bunce-2020-xcultural-competence`. The formulation of the statistical model above produces divergent transitions in Rstan (Stan Development Team 2017), and I have been unable to devise a formulation that eliminates them. This means that the Hamiltonian Monte Carlo sampling algorithm is unable to thoroughly explore the posterior distribution, potentially leading to biased posterior estimates of the parameters of interest. As explained below, this bias does not appear to be severe. However, results of this analysis must be viewed as preliminary, and potentially inaccurate.

With this caveat in mind, Figures A.18 and A.19 show posterior probability distributions of the model parameters after 5 and 25 (respectively) recursions of the theoretical model. Black distributions were fit to all Matsigenka and Mestizos, while blue distributions were fit only to Matsigenka educated among Mestizos and Mestizo employers of Matsigenka (i.e., individuals with more inter-group interaction experience). Note that, regardless of the number of recursion steps, estimated values of the a parameters (in-group affinity) are lower for Matsigenka and Mestizos with more inter-group interaction experience. This coincides with the effect shown in Figure 1, and lends further support to the interpretation that the theoretical model represents processes relevant in the real world. Figures A.20 and A.21 use the means of these posterior estimates to parameterize the full model in the main text, with 5 and 25 recursion steps, respectively. At the appropriate time steps, model predictions approximately coincide with the observed frequencies used to estimate the model parameters (green bars in Figure 1A). This suggests that, despite the divergent transitions during sampling of the statistical model in Equations A.33 - A.43, bias in the posterior parameter estimates is not very severe.

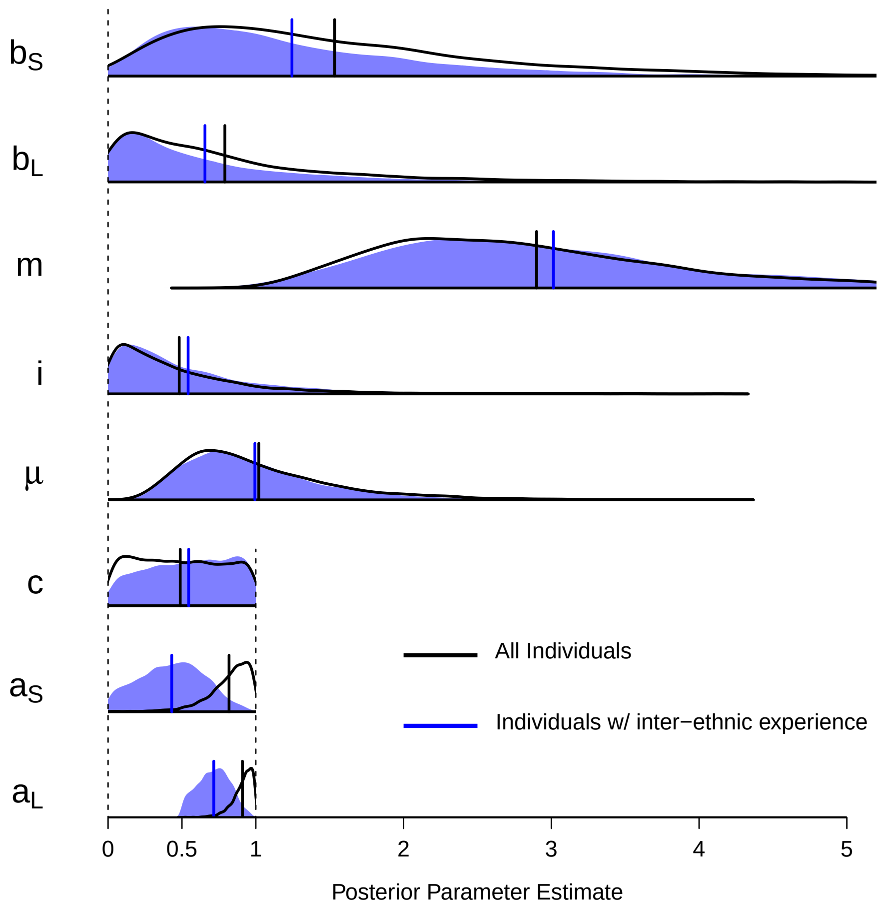

FIGURE A.18. Posterior probability estimates for parameters in the theoretical model after five recursion steps, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$ and $p_{S22} = p_{L11} = 0.1$. Black: estimation using observed phenotype frequencies among all Matsigenka (Group S) and Mestizos (Group L). Blue: estimation using observed phenotype frequencies among only Matsigenka with inter-ethnic education experience and Mestizos with inter-ethnic employer experience. Vertical lines are the means of each posterior distribution. Note that estimated values of the a parameters (in-group affinity) are lower for Matsigenka and Mestizos with more intergroup interaction experience. Compare this effect, and the estimated parameter values, to those of the model in Figure 1. As a warning, estimation produced divergent transitions, so the distributions shown here are potentially biased in unpredictable ways.

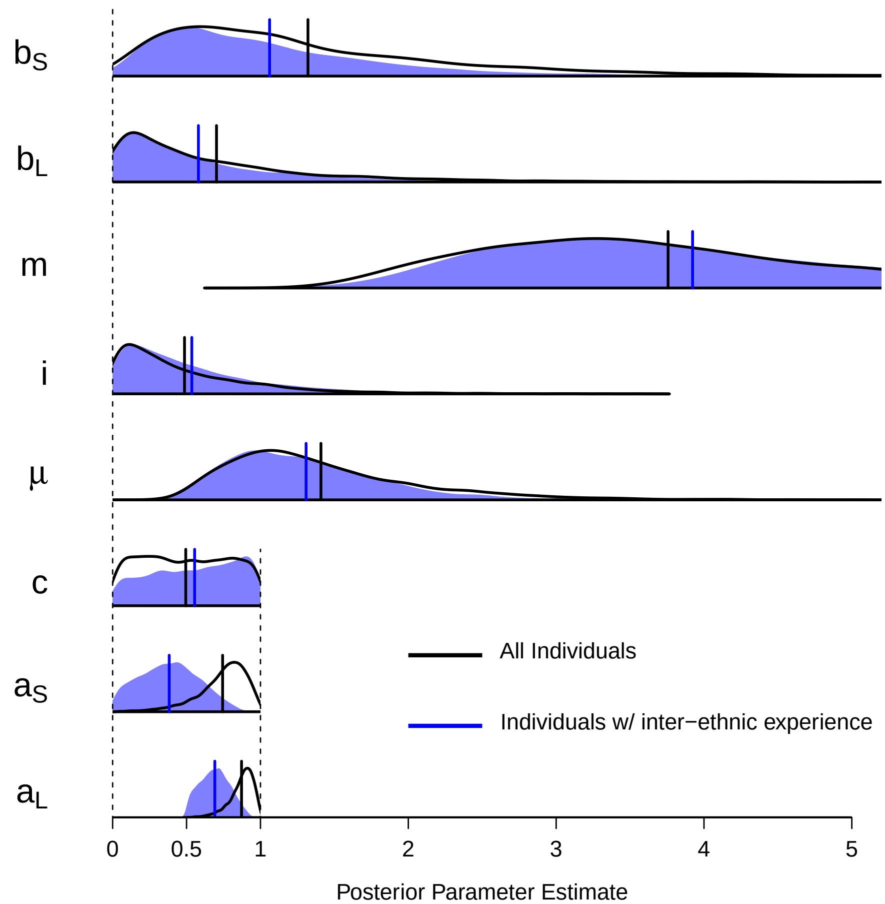

FIGURE A.19. Posterior probability estimates for parameters in the theoretical model after 25 recursion steps. Interpretation is analogous to Figure A.18. Comparing with the previous figure, note that the primary effect of increasing the number of recursion steps is to the increase the estimated values of m, the learning cost of cross-cultural competence, and μ, the payoff bias in norm adoption decisions.

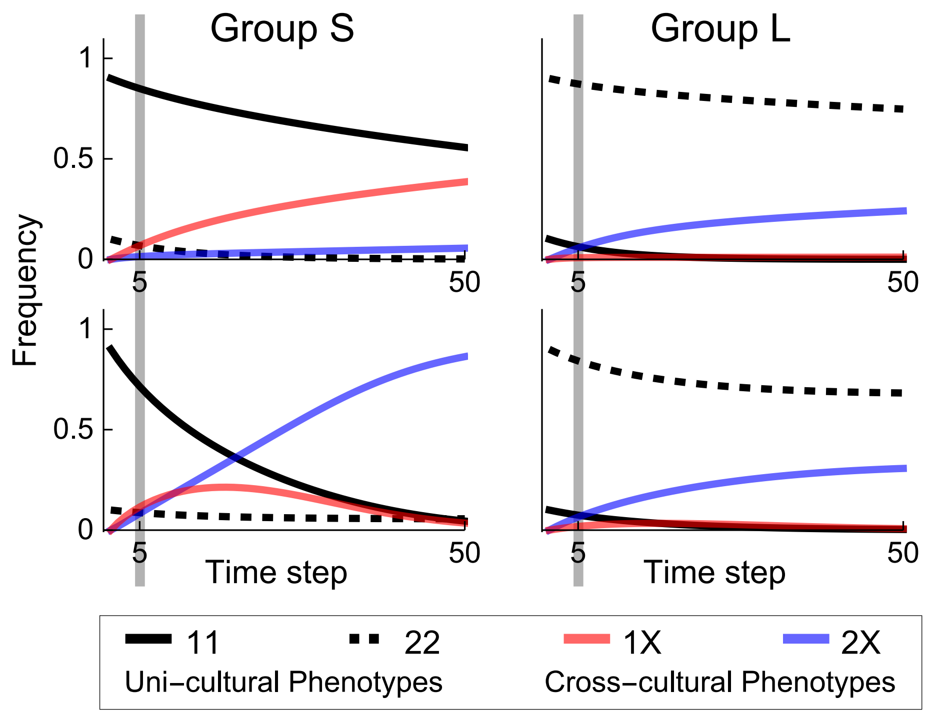

Figure A.20. Phenotype frequency trajectories for Group S (Matsigenka) and Group L (Mestizos) simulated from the full model with initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$, $p_{S22} = p_{L11} = 0.1$, and the values of other model parameters assigned means of the posterior probability distributions from the statistical model in Equations A.33 - A.43, using five recursion steps of the theoretical model ($tmax = 5$). Upper row - parameters estimated using data from all Matsigenka and Mestizos: $(b_S, b_L) = (1.53, 0.79)$, $c = 0.49$, $\mu = 1.02$, $m = 2.9$, $i = 0.48$, and $(a_S, a_L) = (0.82, 0.91)$. Lower row - parameters estimated using data from Matsigenka and Mestizos with more inter-ethnic interaction experience: $(b_S, b_L) = (1.24, 0.66)$, $c = 0.55$, $\mu = 0.99$, $m = 3.01$, $i = 0.54$, and $(a_S, a_L) = (0.43, 0.72)$. Compare the predicted phenotype frequencies at time step 5 to the observed phenotype frequencies (green bars) in the top row of Figure 1A.

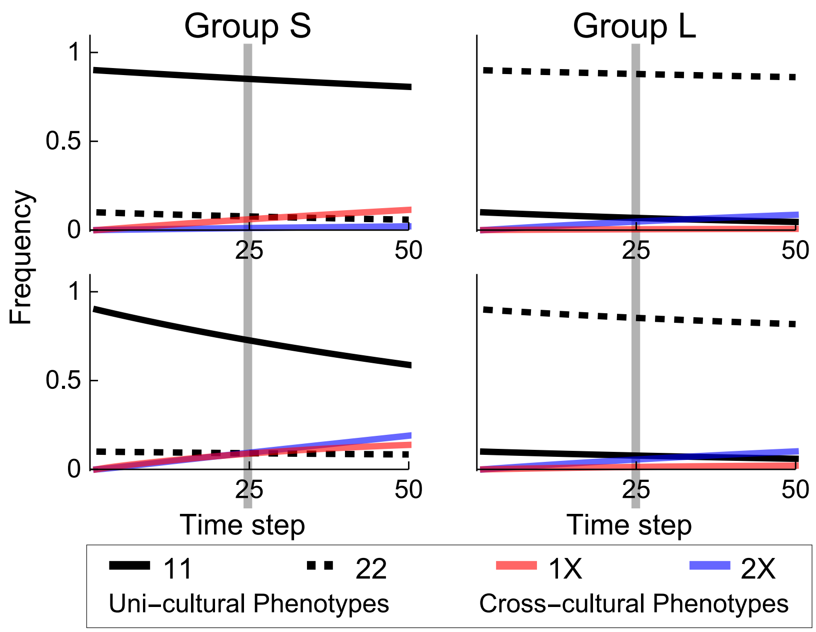

FIGURE A.21. Phenotype frequency trajectories for Group S (Matsigenka) and Group L (Mestizos) simulated from the full model with initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$, $p_{S22} = p_{L11} = 0.1$, and the values of other model parameters assigned means of the posterior probability distributions from the statistical model in Equations A.33 - A.43, using 25 recursion steps of the theoretical model ($tmax = 25$). Upper row - parameters estimated using data from all Matsigenka and Mestizos: $(b_S, b_L) = (1.32, 0.70)$, $c = 0.49$, $\mu = 1.41$, $m = 3.76$, $i = 0.49$, and $(a_S, a_L) = (0.74, 0.87)$. Lower row - parameters estimated using data from Matsigenka and Mestizos with more inter-ethnic interaction experience: $(b_S, b_L) = (1.06, 0.58)$, $c = 0.55$, $\mu = 1.31$, $m = 3.92$, $i = 0.54$, and $(a_S, a_L) = (0.38, 0.69)$. Compare the predicted phenotype frequencies at time step 25 to the observed phenotype frequencies (green bars) in the top row of Figure 1A.

B.4.2. *Interpreting Model Time Steps.* Each time step in the theoretical model is an opportunity for everyone in the population to engage in one coordination attempt, and, potentially, to adopt a different phenotype. The relationship between changes in phenotype frequencies and the time steps (i.e., the rate of phenotype evolution) is determined by the values of the model's parameters, and, in particular, the cost of norm learning, m. Increasing m decreases the rate of phenotype evolution, as it decreases the advantage of CCC phenotypes over UCC phenotypes, and thereby slows the rate of transition from UCC to CCC phenotypes. This can be seen, for instance, in Figures A.11 and A.12, where, when m is high, it takes considerably longer for $S11$ to go to extinction (i.e., all $S11$ individuals transition to a CCC phenotype) in group S, especially when inter-group interaction is low (high a_S). Thus, by adjusting the value of m, empirical estimates of phenotype frequencies in a population at a single point in time (such as the estimates presented in the main text) can be made to correspond to any arbitrary time step in the model. This is demonstrated by comparing the mean posterior estimates for the parameter m in Figures A.18 and A.19, where model parameter values are estimated assuming, arbitrarily, that the empirical data correspond to model predictions at 5 and 25 time steps, respectively. If the empirical data are assumed to correspond to model predictions at 25 time steps (Figure A.19), the estimated mean value of m must be increased (to approximately 4) relative to its estimated mean value (approximately 3) when the data are assumed to correspond to model predictions at 5 time steps (Figure A.18).

In the main text, the empirically-measured phenotype frequencies among Matsigenka and Mestizos are shown to correspond to model predictions at approximately 5 of the model's time steps, given an arbitrarily-chosen value of 1 for the parameter m (Figure 1). Assuming a different (e.g., larger) value of m would result in a slowing of phenotype evolution in the model, such that the empirically-measured phenotype frequencies would correspond to model predictions at a different (i.e., later) time step. For this reason, it cannot be asserted that this analysis provides evidence that Matsigenka and Mestizos have attempted only five coordination interactions in the domain of fairness. Rather, the argument for the plausibility of the model, given the data, relies on the correspondence between model and data with regard to the effect of increasing the (presumed, based on particular inter-ethnic experience) frequency of inter-group interaction. A rigorous test of this model's predictions requires measurement of phenotype frequencies in a population at several (i.e., at least two) distinct points in time. Using the Bayesian procedure above to fit the model to such longitudinal data will allow estimates of the actual value of m, along with values of all of the other parameters in the model. A test of the explanatory power of the model would then entail comparing the fit of this model's predictions against the fit of an alternative model's predictions (e.g., using the WAIC metric: McElreath (2020)), in conjunction with an examination of estimated parameter values of the fitted models in order to judge such values' ethnographic plausibility.

Thus, interpretation of the meaning of time steps in the theoretical model is intimately tied to the assumed values of the model's parameters, particularly m. One time step is equal to one attempted coordination interaction and one opportunity to change one's phenotype, conditional on the assumed cost m of learning a new norm (and on the assumed values of all of the other model parameters). A further complication of attempting to apply this simple model directly to real-world data is the model's assumption that the probabilities of intra- and inter-group interaction are constant through time, and the same for every individual in the population, i.e., one time step has the same meaning for every individual over an extended period of their lives. This is unlikely to be realistic for most domains of interaction, as some individuals have (or seek) more opportunities for interaction than other individuals. However, certain ethnographic contexts may provide reasonable approximations of this assumption. For instance, based on my experience working in Matsigenka-Mestizo inter-ethnic boarding schools, the probability of interaction between teachers (of one ethnicity) and students (of a different ethnicity) in the domain of pedagogical norms may be approximately constant over time, and approximately equal for all students and teachers.

B.5. **Model without Cross-cultural Competence.**

B.5.1. *Model Design.* Many models of norm dynamics in group-structured populations do not include cross-cultural competence, yet can produce stable mixed equilibria, thereby preserving cultural norm diversity. However, as explained in Appendix A, at such mixed equilibria, individuals with different norms cannot coordinate (correlatively) with each other. These equilibria are therefore undesirable from the perspective of an integrative society, as defined in the main text. In contrast, by incorporating CCC, the model in the main text can produce stable mixed equilibria in which all individuals can coordinate. This new model is developed using a set of assumptions (apart from CCC) that differ from those of many previous models. It is therefore of interest to know whether this model, in the absence of CCC, can produce stable mixed equilibria comparable to those produced by other models. Here I show this to be the case.

I construct a model variant that contains only UCC phenotypes (i.e., individuals know and prefer only a single norm). An individual who miscoordinates in a given time step may change her phenotype (i.e., her norm) to that which she perceives will yield the highest mean payoff in the next time step. I retain assumptions of group-level differences in size and power from the model in the main text, as well as the concept of identity valuation (i), allowing this to contribute to the payoffs perceived to be associated with a given phenotype. These assumptions are represented in the interaction Table A.6, from which the recursions in Equations A.44 and A.45 are derived.

$$\begin{aligned}
p'_{S11} = {} & a_S p_{S11} \left\{ p_{S11} + p_{S22} \text{logit}^{-1} \left[\mu(\tilde{w}_{S11} - \tilde{w}_{S22}) \right] \right\} + \\
& (1 - a_S) p_{S11} \left\{ p_{L11} + p_{L22} \text{logit}^{-1} \left[\mu(\tilde{w}_{S11} - \tilde{w}_{S22}) \right] \right\} + \\
& a_S p_{S22} p_{S11} \text{logit}^{-1} \left[\mu(\tilde{w}_{S11} - \tilde{w}_{S22}) \right] + \\
& (1 - a_S) p_{S22} p_{L11} \text{logit}^{-1} \left[\mu(\tilde{w}_{S11} - \tilde{w}_{S22}) \right] \qquad \text{(A.44)}
\end{aligned}$$

$$p'_{S22} = 1 - p'_{S11} \qquad \text{(A.45)}$$

p'_{L22} and p'_{L11} are obtained by reversing all group and norm subscripts in Equations A.44 and A.45, respectively. Perceived phenotype-specific average payoffs $\tilde{w}_{S11}$ and $\tilde{w}_{S22}$ are defined as in Equations 1 - 3 and 6, with $p_{S1X} = p_{S2X} = 0$.

TABLE A.6. Interaction table for individuals with phenotypes $S11$ and $S22$, assuming no cross-cultural competence

Self	Other	Pr(Self, Other)[a]	Pr($S11$)[b]	Pr($S22$)
S11	S11	$a_S p_{S11}^2$	1	0
S11	S22	$a_S p_{S11} p_{S22}$	$\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S22})]$	$\text{logit}^{-1}[\mu(\tilde{w}_{S22} - \tilde{w}_{S11})]$
S11	L11	$(1 - a_S) p_{S11} p_{L11}$	1	0
S11	L22	$(1 - a_S) p_{S11} p_{L22}$	$\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S22})]$	$\text{logit}^{-1}[\mu(\tilde{w}_{S22} - \tilde{w}_{S11})]$
S22	S11	$a_S p_{S22} p_{S11}$	$\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S22})]$	$\text{logit}^{-1}[\mu(\tilde{w}_{S22} - \tilde{w}_{S11})]$
S22	S22	$a_S p_{S22}^2$	0	1
S22	L11	$(1 - a_S) p_{S22} p_{L11}$	$\text{logit}^{-1}[\mu(\tilde{w}_{S11} - \tilde{w}_{S22})]$	$\text{logit}^{-1}[\mu(\tilde{w}_{S22} - \tilde{w}_{S11})]$
S22	L22	$(1 - a_S) p_{S22} p_{L22}$	0	1

[a]Pr(interaction between Self phenotype and Other phenotype | level of in-group affinity a_S)
[b]Pr(Self=S11 after updating | current phenotype of Self, current phenotype of Other, a_S). Column 5 interpreted analogously.

B.5.2. *Model Analysis.* As shown in Figure A.22, in the absence of CCC, high identity valuation (i) can result in a stable mixed equilibrium, even if inter-group interaction is frequent, group S is a minority, and S members have low bargaining power. To retain a high frequency of norm 1 in group S, the value of i must be sufficient to outweigh the perceived mean cost to S members of miscoordinating with L members, most of whom have norm 2. Note that, at this mixed equilibrium (bottom row of Figure A.22), more than half of S members' interactions occur with the out-group ($a_S = 0.4$). Most of these S individuals have norm 1, while most L individuals have norm 2. Thus, miscoordination is extremely common at this equilibrium, and it is therefore a poor representation of an integrative society (in common with the equilibria of most other models discussed in Appendix A). Furthermore, in the real world, such a situation is unlikely to be stable, as people would be expected to use any means available (e.g., group identity markers that covary with norms: McElreath et al. (2003)) to avoid such unprofitable interactions. As shown in the main text, one ethnographically-plausible way to achieve a mixed equilibrium (sustained cultural diversity) with universal coordination is to take advantage of the fact that people are capable of becoming cross-culturally competent.

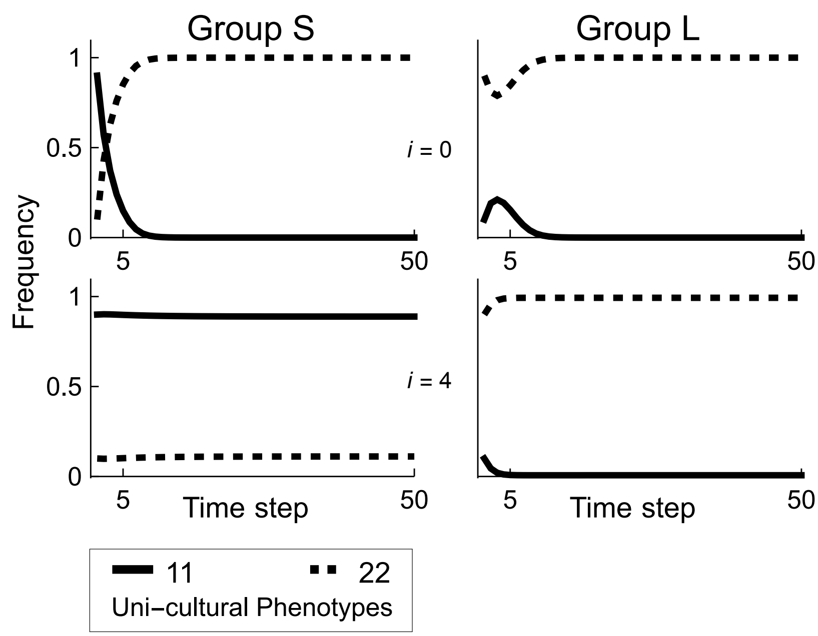

FIGURE A.22. Model simulations of phenotype frequency trajectories for the low-power minority group (S) and high-power majority group (L) simulated from a model with no cross-cultural competence, where: $(b_S, b_L) = (1, 0)$, $(a_S, a_L) = (0.4, 0.7)$, $\mu = 1$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$, $p_{S22} = p_{L11} = 0.1$, given $i = 0$ (top row) and $i = 4$ (bottom row). Compare to Figures 2A and A.23.

B.6. **Model with Forgetting.**

B.6.1. *Model Design.* Cross-cultural competence, by definition, requires individuals to retain in their memory knowledge of a norm that they do not personally prefer. As a simplifying assumption in the model in the main text, individuals can never forget a norm that they learn at some point in their lifetime, making CCC an absorbing state. Here I examine the consequences of this assumption by modifying the original model to include an extreme form of forgetting: UCC individuals who make a payoff-biased decision to learn a new norm (and thereby become one of the two forms of CCC) immediately forget whichever of the two norms they know but do not prefer. Therefore, they immediately return to one of the two forms of UCC – but not necessarily the same phenotype with which they started off. I retain the assumption in the original model that UCC individuals preferring a given norm must pass through a CCC phenotype before transitioning to another UCC phenotype preferring a different norm (Figure A.8). However, now such transitions occur between time steps, resulting in a situation in which CCC phenotypes are never actually observed in the population. These assumptions are represented in the interaction Table A.7, from which the recursions in Equations A.46 and A.47 are derived.

$$
\begin{aligned}
p'_{S11} = {} & a_S p_{S11} \Bigg(p_{S11} + \frac{p_{S22}}{P_{S11\text{tot}}} \left\{ \text{logit}^{-1}\left[\mu(\tilde{w}_{S11} - \tilde{w}_{S1X})\right] \text{logit}^{-1}\left[\mu(\tilde{w}_{S11} - \tilde{w}_{S2X})\right] + \right. \\
& \left. \text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S11})\right] \text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})\right] \right\} \Bigg) + \\
& (1-a_S) p_{S11} \Bigg(p_{L11} + \frac{p_{L22}}{P_{S11\text{tot}}} \left\{ \text{logit}^{-1}\left[\mu(\tilde{w}_{S11} - \tilde{w}_{S1X})\right] \text{logit}^{-1}\left[\mu(\tilde{w}_{S11} - \tilde{w}_{S2X})\right] + \right. \\
& \left. \text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S11})\right] \text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})\right] \right\} \Bigg) + \\
& a_S p_{S22} p_{S11} \frac{\text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S22})\right] \text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})\right]}{P_{S22\text{tot}}} + \\
& (1-a_S) p_{S22} p_{L11} \frac{\text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S22})\right] \text{logit}^{-1}\left[\mu(\tilde{w}_{S1X} - \tilde{w}_{S2X})\right]}{P_{S22\text{tot}}}
\end{aligned}
\tag{A.46}
$$

$$
p'_{S22} = 1 - p'_{S11} \tag{A.47}
$$

p'_{L22} and p'_{L11} are obtained by reversing all group and norm subscripts in Equations A.46 and A.47, respectively. Perceived phenotype-specific average payoffs ($\tilde{w}$) are defined in Equations 1 - 6.

Table A.7. Interaction table for individuals with phenotypes $S11$ and $S22$, assuming instantaneous forgetting of any non-preferred norm

Self	Other	Pr(Self, Other)[a]	Pr($S11$)[b]	Pr($S22$)
S11	S11	$a_S p_{S11}^2$	1	0
S11	S22	$a_S p_{S11} p_{S22}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S1X})]\text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S2X})]+\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S11})]\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S11\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S11})]\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S11\text{tot}}}$
S11	L11	$(1-a_S) p_{S11} p_{L11}$	1	0
S11	L22	$(1-a_S) p_{S11} p_{L22}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S1X})]\text{logit}^{-1}[\mu(\tilde{w}_{S11}-\tilde{w}_{S2X})]+\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S11})]\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S11\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S11})]\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]}{P_{S11\text{tot}}}$
S22	S11	$a_S p_{S22} p_{S11}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S22})]\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S22\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S22})]\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]+\text{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S1X})]\text{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S2X})]}{P_{S22\text{tot}}}$
S22	S22	$a_S p_{S22}^2$	0	1
S22	L11	$(1-a_S) p_{S22} p_{L11}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S22})]\text{logit}^{-1}[\mu(\tilde{w}_{S1X}-\tilde{w}_{S2X})]}{P_{S22\text{tot}}}$	$\frac{\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S22})]\text{logit}^{-1}[\mu(\tilde{w}_{S2X}-\tilde{w}_{S1X})]+\text{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S1X})]\text{logit}^{-1}[\mu(\tilde{w}_{S22}-\tilde{w}_{S2X})]}{P_{S22\text{tot}}}$
S22	L22	$(1-a_S) p_{S22} p_{L22}$	0	1

[a]Pr(interaction between Self phenotype and Other phenotype | level of in-group affinity a_S)

[b]Pr(Self=S11 after updating | current phenotype of Self, current phenotype of Other, a_S). Column 5 interpreted analogously. $P_{S11\text{tot}}$ and $P_{S22\text{tot}}$ are defined in Table A.3.

B.6.2. *Model Analysis.* This assumption of extreme forgetting makes it even easier (i.e., requires a lower value of i) to sustain minority-typical norm 1 at equilibrium (compare bottom rows of Figures A.23 and 2A). Here, all $S1X$ individuals change immediately into $S11$ individuals, which therefore maintain a higher frequency in group S, as long as i is perceived as sufficient to offset the cognitive dissonance cost of inter-group coordination, c, suffered by $S1X$ individuals using their non-preferred norm. $S2X$ individuals change immediately into $S22$. For most in-group interactions, because of the higher frequency of $S11$, $S2X$ phenotypes would be perceived to suffer c and $S22$ phenotypes would be perceived to suffer miscoordination. Thus, even if inter-group interaction is frequent (low a_S), as long as i is sufficiently high, $S22$ will be held at a low frequency in group S (Figure A.23, bottom row).

As shown, with moderate values of i, mixed equilibria are possible even with extreme degrees of non-preferred norm forgetfulness. Importantly, however, at such an equilibrium, most inter-group interactions would result in miscoordination (analogous to a model without CCC: Appendix B.5), an undesirable outcome for an integrative society (as defined in the main text). The extreme forgetfulness in this variant of the model seems very unnatural, but it suggests that the original model's assumption of no forgetting is not crucial for arriving at a mixed equilibrium (i.e., sustaining cultural diversity). However, it also suggests that forgetting non-preferred norms within one's lifetime, by degrading CCC, works against the formation of an integrative society in which cultural diversity is sustained and everyone can coordinate with everyone else in a given domain.

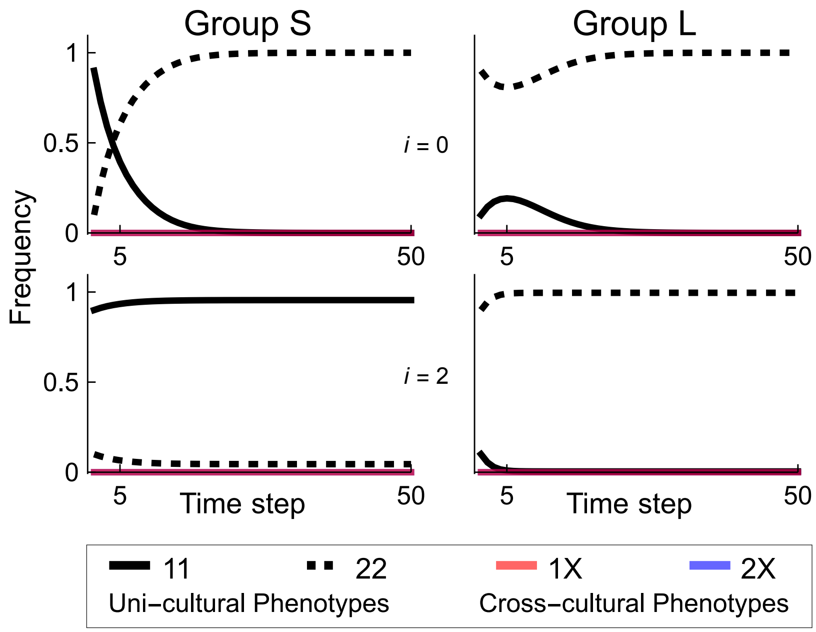

FIGURE A.23. Model simulations of phenotype frequency trajectories for the low-power minority group (S) and high-power majority group (L) simulated from a model with instantaneous forgetting of the non-preferred norm, where: $(b_S, b_L) = (1, 0)$, $(a_S, a_L) = (0.4, 0.7)$, $c = 0.1$, $\mu = m = 1$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$, $p_{S22} = p_{L11} = 0.1$, given $i = 0$ (top row) and $i = 2$ (bottom row). Compare to Figures 2A and A.22.

B.7. **Model with Traditional Replication Assumptions.**

B.7.1. *Model Design.* Here I modify the original full model by changing several assumptions in order to more closely match those commonly used in cultural evolutionary game theory. During the interaction phase of the model, individuals receive payoffs as a result of their interactions (Table A.2) and the utility they receive from their sense of cultural identity. As in the original model, this identity-based utility is the product of: 1) the degree to which they value a sense of belonging to their in-group, i; 2) the frequency of their preferred norm among their fellow in-group members; and 3) the frequency of their non-preferred norm among out-group members. Note that here, in contrast to the original model, people know the actual frequencies of the norms preferred by in- and out-group members (or such frequencies in a representative sample). Payoffs received by phenotypes in group S during the interaction phase are given in Equations A.48 - A.51.

$$\begin{aligned} w_{S11} = & a_S(p_{S11} + p_{S1X} + p_{S2X}) + \\ & (1 - a_S)\left[p_{L11}(1 + b_S) + p_{L1X}(1 + b_S) + p_{L2X}(1 + b_S)\right] + \\ & i(p_{S11} + p_{S1X})(p_{L22} + p_{L2X}) \end{aligned} \tag{A.48}$$

$$\begin{aligned} w_{S1X} = & a_S\left[p_{S11} + p_{S1X} + p_{S2X}(1 - \tfrac{1}{2}c) + p_{S22}(1 - c)\right] + \\ & (1 - a_S)\left[p_{L11}(1 + b_S) + p_{L1X}(1 + b_S) + \right. \\ & \left. p_{L2X}(1 + b_S - \tfrac{1}{2}c) + p_{L22}(1 + b_S - c)\right] - m + \\ & i(p_{S11} + p_{S1X})(p_{L22} + p_{L2X}) \end{aligned} \tag{A.49}$$

$$\begin{aligned} w_{S2X} = & a_S\left[p_{S11}(1 - c) + p_{S1X}(1 - \tfrac{1}{2}c) + p_{S2X} + p_{S22}\right] + \\ & (1 - a_S)\left[p_{L11}(1 + b_S - c) + p_{L1X}(1 + b_S - \tfrac{1}{2}c) + \right. \\ & \left. p_{L2X}(1 + b_S) + p_{L22}(1 + b_S)\right] - m + \\ & i(p_{S22} + p_{S2X})(p_{L11} + p_{L1X}) \end{aligned} \tag{A.50}$$

$$\begin{aligned} w_{S22} = & a_S(p_{S1X} + p_{S2X} + p_{S22}) + \\ & (1 - a_S)\left[p_{L1X}(1 + b_S) + p_{L2X}(1 + b_S) + p_{L22}(1 + b_S)\right] + \\ & i(p_{S22} + p_{S2X})(p_{L11} + p_{L1X}) \end{aligned} \tag{A.51}$$

Payoffs to phenotypes in group L are obtained by reversing all group subscripts in these equations. Compare Equations A.48 - A.51 to Equations A.13 - A.16 and A.19 - A.22.

In the updating phase, individuals choose an in-group member at random with whom to compare payoffs, and from whom to potentially copy a phenotype (as in: McElreath and Boyd 2007; Boyd and Richerson 1985). I assume that individuals can observe and compare payoffs with, and copy phenotypes from, only in-group members (co-ethnics). This assumption derives from the fact that members of different ethnic groups may display wealth (a possible result of interaction payoffs) in different ways, which are not always apparent to members of other groups. For instance, given a sum of money, minority indigenous Matsigenka may invest in buying a boat motor in order better access food resources along the river, while majority-culture Mestizos might invest in the education of their children who are living and studying in a distant city. In such a situation, it may be more difficult for Matsigenka (who are less familiar with such educational costs) to compare their payoffs against those of Mestizos, than it is for them to compare their payoffs against those of other Matsigenka. As a result of comparing her payoff with that of a randomly-chosen in-group member, an individual may attempt to adopt the phenotype of this fellow in-group member. As in the original model, the adoption decision is modelled as a logistic function of the difference in average payoffs between phenotypes, where the parameter μ determines the degree of bias toward adopting (or retaining) the phenotype with the higher average payoff. Note that these replication assumptions differ from those of the original model in that individuals make decisions to change (or not) their phenotype based on randomly choosing, and comparing themselves to, an in-group member whose phenotype and payoffs they can directly observe, rather than basing such decisions on their (potentially inaccurate) perceptions of how their personal payoffs are likely to change in the future if they change their phenotype.

As in the original model, I assume that norms cannot be unlearned within an individual's lifetime. Thus, it may not always be possible to adopt the phenotype of another individual, even if this individual received a higher payoff. When a decision is made to adopt a new phenotype, I assume that individuals attempt to adopt a phenotype that is as close as possible (given their current phenotype) to that of the individual whose phenotype they are attempting to copy. For instance, assume a focal individual with phenotype $S11$ is attempting to copy the phenotype of an $S22$ individual. The focal cannot unlearn norm 1, so she will change her phenotype to $S2X$, such that she favors norm 2 (like the individual she is attempting to copy) but still retains knowledge of norm 1 (i.e., she becomes cross-culturally competent). These assumptions are represented in the interaction Table A.8, from which the recursions in Equations A.52 and A.53 are derived.

$$\begin{aligned} p'_{S11} = \; & p_{S11} \{ p_{S11} + p_{S1X} \text{logit}^{-1} [\mu(w_{S11} - w_{S1X})] + \\ & p_{S2X} \text{logit}^{-1} [\mu(w_{S11} - w_{S2X})] + \\ & p_{S22} \text{logit}^{-1} [\mu(w_{S11} - w_{S22})] \} \end{aligned} \tag{A.52}$$

$$\begin{aligned} p'_{S1X} = \; & p_{S11} p_{S1X} \text{logit}^{-1} [\mu(w_{S1X} - w_{S11})] + \\ & p_{S1X} (p_{S11} + p_{S1X}) + \\ & p_{S1X} p_{S2X} \left\{ \frac{1}{2} + \frac{1}{2} \text{logit}^{-1} [\mu(w_{S1X} - w_{S2X})] \right\} + \\ & p_{S1X} p_{S22} \text{logit}^{-1} [\mu(w_{S1X} - w_{S22})] + \\ & p_{S2X} p_{S11} \text{logit}^{-1} [\mu(w_{S11} - w_{S2X})] + \\ & p_{S2X} p_{S1X} \frac{1}{2} \text{logit}^{-1} [\mu(w_{S1X} - w_{S2X})] + \\ & p_{S22} p_{S11} \text{logit}^{-1} [\mu(w_{S11} - w_{S22})] + \\ & p_{S22} p_{S1X} \text{logit}^{-1} [\mu(w_{S1X} - w_{S22})] \end{aligned} \tag{A.53}$$

p'_{S2X} and p'_{S22} are obtained by reversing all norm subscripts in Equations A.53 and A.52, respectively.

Table A.8. Interaction table for individuals with phenotypes $S11$, $S1X$, $S2X$, and $S22$, under traditional replication assumptions

Self	Other	Pr(Self, Other)[a]	Pr($S11$)[b]	Pr($S1X$)	Pr($S2X$)	Pr($S22$)
S11	S11	p_{S11}^2	1	0	0	0
S11	S1X	$p_{S11}p_{S1X}$	$\text{logit}^{-1}[\mu(w_{S11}-w_{S1X})]$	$\text{logit}^{-1}[\mu(w_{S1X}-w_{S11})]$	0	0
S11	S2X	$p_{S11}p_{S2X}$	$\text{logit}^{-1}[\mu(w_{S11}-w_{S2X})]$	0	$\text{logit}^{-1}[\mu(w_{S2X}-w_{S11})]$	0
S11	S22	$p_{S11}p_{S22}$	$\text{logit}^{-1}[\mu(w_{S11}-w_{S22})]$	0	$\text{logit}^{-1}[\mu(w_{S22}-w_{S11})]$	0
S1X	S11	$p_{S1X}p_{S11}$	0	1	0	0
S1X	S1X	p_{S1X}^2	0	1	0	0
S1X	S2X	$p_{S1X}p_{S2X}$	0	$\frac{1}{2}+\frac{1}{2}\text{logit}^{-1}[\mu(w_{S1X}-w_{S2X})]$	$\frac{1}{2}\text{logit}^{-1}[\mu(w_{S2X}-w_{S1X})]$	0
S1X	S22	$p_{S1X}p_{S22}$	0	$\text{logit}^{-1}[\mu(w_{S1X}-w_{S22})]$	$\text{logit}^{-1}[\mu(w_{S22}-w_{S1X})]$	0
S2X	S11	$p_{S2X}p_{S11}$	0	$\text{logit}^{-1}[\mu(w_{S11}-w_{S2X})]$	$\text{logit}^{-1}[\mu(w_{S2X}-w_{S11})]$	0
S2X	S1X	$p_{S2X}p_{S1X}$	0	$\frac{1}{2}\text{logit}^{-1}[\mu(w_{S1X}-w_{S2X})]$	$\frac{1}{2}+\frac{1}{2}\text{logit}^{-1}[\mu(w_{S2X}-w_{S1X})]$	0
S2X	S2X	p_{S2X}^2	0	0	1	0
S2X	S22	$p_{S2X}p_{S22}$	0	0	1	0
S22	S11	$p_{S22}p_{S11}$	0	$\text{logit}^{-1}[\mu(w_{S11}-w_{S22})]$	0	$\text{logit}^{-1}[\mu(w_{S22}-w_{S11})]$
S22	S1X	$p_{S22}p_{S1X}$	0	$\text{logit}^{-1}[\mu(w_{S1X}-w_{S22})]$	0	$\text{logit}^{-1}[\mu(w_{S22}-w_{S1X})]$
S22	S2X	$p_{S22}p_{S2X}$	0	0	$\text{logit}^{-1}[\mu(w_{S2X}-w_{S22})]$	$\text{logit}^{-1}[\mu(w_{S22}-w_{S2X})]$
S22	S22	p_{S22}^2	0	0	0	1

[a]Pr(interaction between Self phenotype and Other phenotype)

[b]Pr(Self=S11 after updating | current phenotype of Self and current phenotype of Other). Columns 5 through 7 interpreted analogously.

B.7.2. *Model Analysis.* The dynamics of this model, as shown in Figure A.24, are broadly similar to those of the original model, with a few notable exceptions. In the situation of interest, such that members of a minority low-power group (S) engage in frequent inter-ethnic interaction (low a_S) with members of a culturally-distinct majority (L), the top row of Figure A.24 shows that the cross-culturally competent phenotype $S2X$ will reach high frequency in group S when cultural group identity provides no additional utility ($i = 0$). As explained in the main text, if this occurs, S-typical norm 1 is likely to be lost from the population after one generation. However, if group identity is valued ($i > 0$), the cross-cultural phenotype $S1X$ can attain high frequency (bottom row of Figure A.24), and norm 1 can potentially be preserved in group S indefinitely (see main text). Note that the uni-culturally competent phenotype $S11$ is quickly lost from group S. These dynamics are very similar to those in the original model, and demonstrate that the main qualitative results are robust to non-trivial modifications to the assumptions about how individuals decide to adopt new norms.

One major difference between this model and the original is the complete replacement of the uni-culturally competent phenotype ($L22$) by the cross-culturally competent phenotype ($L2X$) in majority group L. In the original model, as explained above, transitions $L22 \rightarrow L2X$ cease as soon as $S11$ is lost from group S, at which point $L22$ never fails to coordinate and average payoffs exceed those to $L2X$. After the disappearance of $S11$, $L22$ individuals never receive less than the maximum possible payoff to coordination, and they therefore do not consider changing their phenotype. In the present model, in contrast, the transitions $L22 \rightarrow L2X$ continue after the disappearance of $S11$. This is because all individuals continue to consider changing their phenotype, even if they receive the maximum possible payoff. For example, during the updating phase of this model, an $L22$ individual may compare herself (at random) with an $L2X$ individual. Even if her payoff is higher than that of the $L2X$ individual, there is a probability less than 0.5 but greater the zero that she will adopt $L2X$. However, because of the assumption that individuals cannot unlearn a norm, all transitions $L2X \rightarrow L22$ are impossible. Thus, as can be seen in Figure A.24, eventually all $L22$ will transition to $L2X$. This has the effect of increasing the average payoff to $S1X$ compared to the original model, where $L22$ is maintained at high frequency in group L. During inter-group interactions, $S1X$ receives a higher average payoff from coordinating with $L2X$ (where only half of interactions entail the cost c) than with $L22$ (where all interactions entail the cost c). Thus, as $L22$ disappears, the average payoff to $S1X$ increases. For this reason, in the present

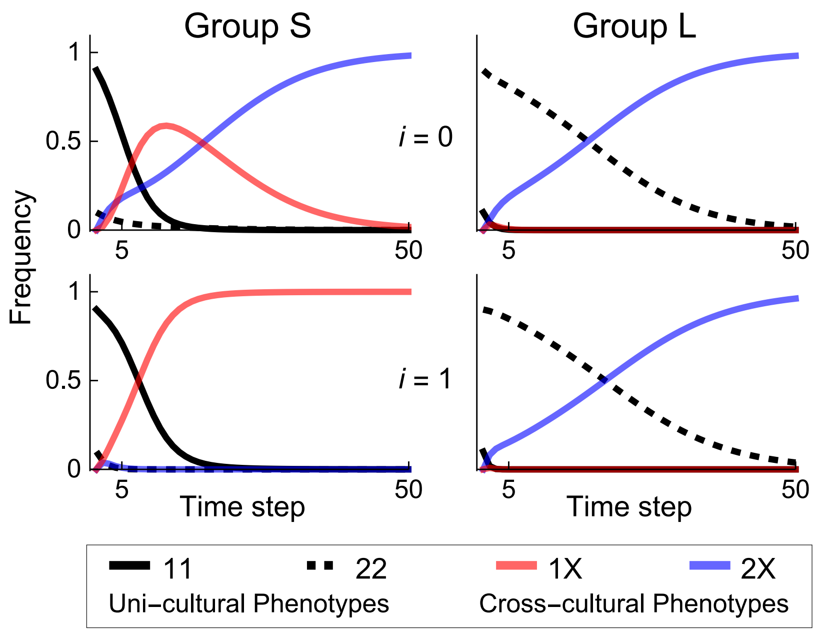

FIGURE A.24. Model simulations of phenotype frequency trajectories for the low-power minority group (S) and high-power majority group (L) simulated from the full model with traditional replication assumptions where: $(b_S, b_L) = (1, 0)$, $(a_S, a_L) = (0.4, 0.7)$, $c = 0.5$, $\mu = 2$, $m = 1$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$, $p_{S22} = p_{L11} = 0.1$, given $i = 0$ (top row) and $i = 1$ (bottom row). Compare to Figure 2A.

model, it is generally easier to sustain $S1X$ than it is in the original model (e.g., it can be done with a lower value of i). Also note that, in contrast to the original model, dynamics in the present model require initial conditions such that both uni-culturally competent phenotypes are present in each group. Thus, in Equation A.53, when $p_{S1X} = 1$ then $p'_{S1X} = p_{S1X}$ and there is no evolution.

A comparison of Figures A.25 and A.26 with Figures A.11 and A.12, reveals that the dynamics of the present model and the original model respond in similar ways to changes in the values of several model parameters, though there are differences worth noting. Figures A.25A and B and A.26A and B, show that it is relatively to easy to sustain $S1X$ in group S when in-group interaction affinity (a) is high. Comparing with the original model, this is because, in the present model, during the few inter-ethnic interactions that do occur, $S1X$ suffers the cognitive dissonance cost c less frequently as a result of the fact that group L is quickly taken over by the cross-culturally competent phenotype $L2X$, as explained above. As inter-group interaction increases (high a), $S2X$ goes to fixation, as it does not suffer the cost c that is suffered by $S1X$ during inter-group interaction (Figure A.25C). When group identity is sufficiently valued (high i), it can more than compensate for this cost c, and $S1X$ can reach fixation (Figure A.25D). This effect is obviously reversed if c is increased (Figures A.25F and A.26D).

Note, that, in contrast to the original model, when i is sufficiently high, $S2X$ can go to extinction in group S, rather than be maintained at low frequency (compare Figures A.24 and 2A). This is because, in the original model, when $S1X$ reaches high frequency, frequent inter-group interactions with $L2X$ and $L22$ result sub-maximal payoffs to $S1X$, and thus a non-zero probability of transitions $S1X \rightarrow S2X$. However, because inter-group interactions are high (low a), $S2X$ interacts often with $L2X$ and $L22$, thereby receiving maximal payoffs, precluding reconsideration of its phenotype, and thereby precluding transitions $S2X \rightarrow S1X$. In the present model, with traditional replication assumptions, $S2X$ always considers changing its phenotype in the updating phase, even when it received maximal payoffs in the interaction phase. Thus, when $S2X$ occurs at low frequency in group S and receives lower average payoffs than $S1X$, the probability of transitions $S2X \rightarrow S1X$ are always high, and $S2X$ goes to extinction.

In contrast to the original model, when inter-group interaction is high (low a), $S2X$ often increases in frequency faster than $S1X$ (compare Figures A.25C-H and A.11C-H). This is because $S11$ individuals consider changing their phenotype in each updating phase to that of a randomly-selected in-group member, by comparing payoffs. Initially, $S11$ individuals observe only $S11$ and $S22$ individuals in the updating phase. The latter observation results in a high probability of a transition $S11 \rightarrow S2X$, because $S22$ has a higher average payoff than $S11$. Similarly, when $S11$ observes $S2X$ in subsequent updating phases, there is a high probability of a transition $S11 \rightarrow S2X$. Initially, $S1X$ is generated only through low-probability transitions $S22 \rightarrow S1X$ resulting from $S22$ individuals observing $S11$ individuals in the updating phase. This low probability is further decreased when payoff-biased phenotype copying (μ) is high (Figures A.25H and A.26F). When i is sufficiently large (relative to b_S and c), the average payoff to $S11$ relative to $S22$ is increased and transitions $S22 \rightarrow S1X$ occur with higher probability, while transitions $S11 \rightarrow S2X$ occur with lower probability (Figures A.24 and A.25D). In the original model, in contrast, $S1X$ is generated with non-zero probability whenever $S11$ interacts with $S22$ or $L22$, thereby receiving a sub-maximal payoff.

In contrast to the original model, increasing the power difference between groups S and L (high b_S) is detrimental to the sustainability of $S1X$ (Figures A.25E and A.26C). This is because high b_S increases the payoff advantage of $S22$ over $S11$ early in the dynamics while $L22$ is still common in group L. This results in a higher probability of $S11 \to S2X$ transitions, and a lower probability of $S22 \to S1X$ transitions when $S11$ and $S22$ individuals observe each other during the updating stage. As noted above, a similar effect is also caused by an increase in the bias toward higher payoffs (μ) during individuals' decisions to modify their phenotype (Figures A.25H and A.26F). High μ increases the probability of $S11 \to S2X$ transitions, and lowers the probability of $S22 \to S1X$ transitions, when $S22$ have higher average payoffs than $S11$.

In contrast to the original model, increasing the learning cost of cross-cultural competence (m) can contribute to the sustainability of $S1X$ (Figures A.25G and A.26E). This is because increasing m decreases the probability of a transition $S11 \to S2X$ when $S11$ and $S2X$ observe each other during the updating phase (but not such transitions resulting from $S11$ and $S22$ observing each other). This maintains a preference for norm 1 at high frequency in group S for a longer time early in the dynamics. When i is high, this, in turn, increases the probability of transitions $S2X \to S1X$, which are unaffected by m.

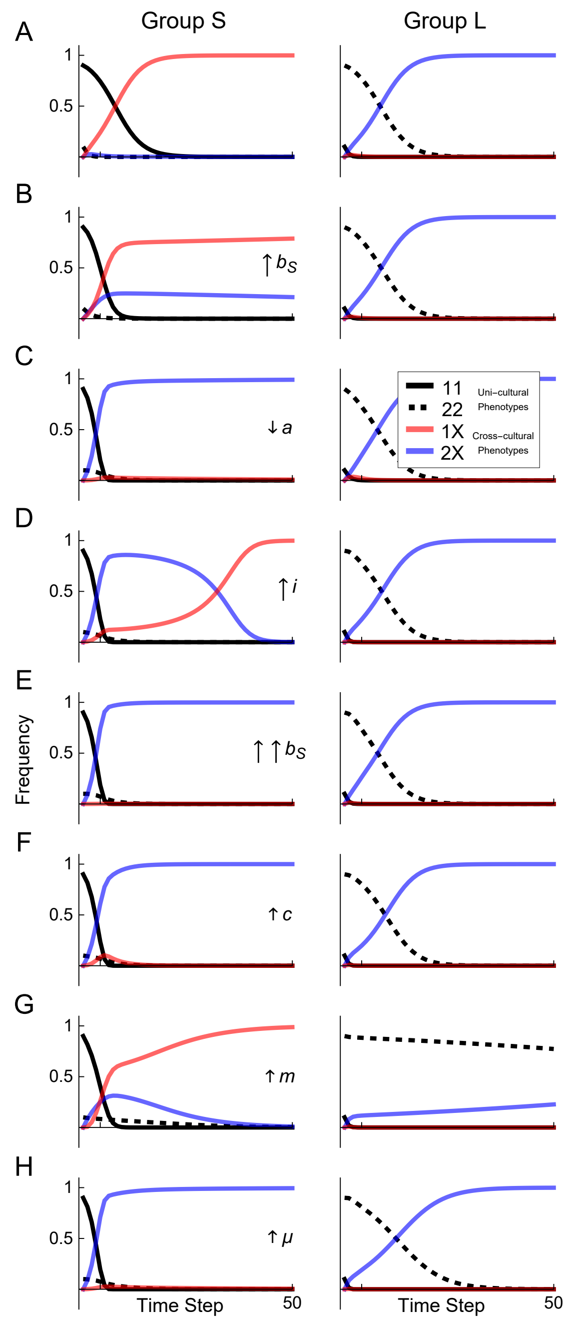

FIGURE A.25. (Caption next page.)

FIGURE A.25. Sensitivity of phenotype trajectories for the minority group (S) and majority group (L), simulating from the full model with traditional replication assumptions. **A)** No power difference between groups, high in-group affinity: $(b_S, b_L) = (0, 0)$, $(a_S, a_L) = (0.8, 0.9)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, $i = 0$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$ and $p_{S22} = p_{L11} = 0.1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (4, 0)$. **C)** Low affinity: sames as **B**, except $(a_S, a_L) = (0.4, 0.7)$. **D)** Identity valuation: sames as **C**, except $i = 1$. **E)** Large power difference: same as **D**, except $(b_S, b_L) = (8, 1)$. **F)** High cognitive dissonance: sames as **D**, except $c = 0.9$. **G)** High cross-cultural learning cost: same as **D**, except $m = 2$. **H)** High payoff-bias for copying: sames as **D**, except $\mu = 3$. Compare to Figure A.11.

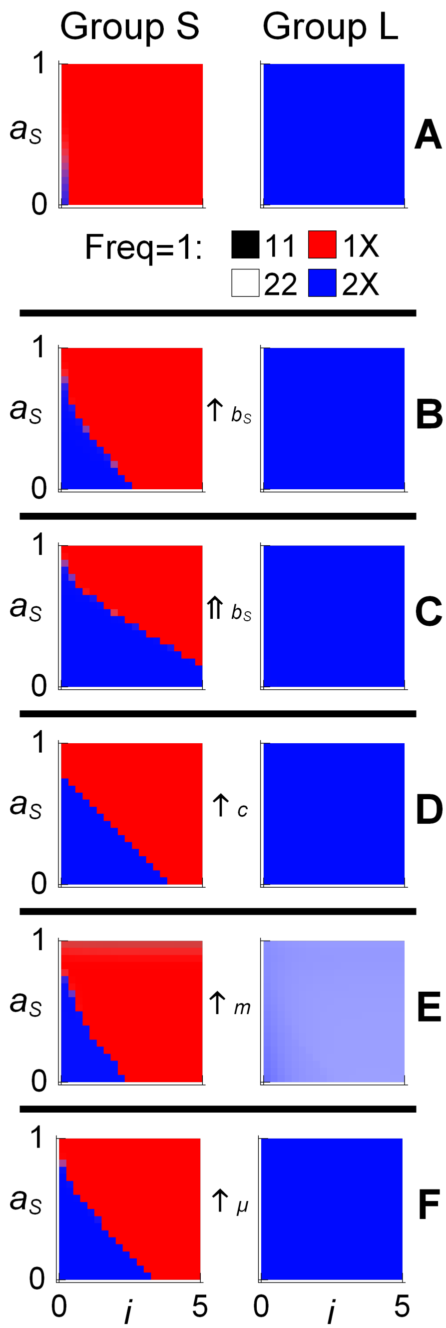

FIGURE A.26. (Caption next page.)

FIGURE A.26. Sensitivity of long-run phenotype frequencies for the minority group (S) and majority group (L), simulating from the full model with traditional replication assumptions for 100 time steps. **A)** No power difference between groups: $(b_S, b_L) = (0, 0)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$ and $p_{S22} = p_{L11} = 0.1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (4, 0)$. **C)** Large power difference: sames as **B**, except $(b_S, b_L) = (8, 1)$. **D)** High cognitive dissonance: sames as **B**, except $c = 0.9$. **E)** High cross-cultural learning cost: sames as **B**, except $m = 2$. **F)** High payoff-bias for copying: same as **B**, except $\mu = 3$. Compare to Figure A.12.

B.8. Model with Non-ergodic Utility.

B.8.1. *Model Design.* Here I modify the original full model by changing the way individuals make decisions about if and how to modify their phenotype. In the original model, individuals make such decisions on the basis of a comparison between the average payoffs that they anticipate each potential phenotype will receive in the next interaction phase, biasing their choice toward that which they perceive will receive the highest average payoff. In the original model, a person's utility in each time step is equated with her interaction payoff, and each payoff is independent of any payoff that she received in past time steps. Thus, a person's utility accumulates over the course of her life by simply adding the payoffs she receives in each time step. In the real world, however, interaction payoffs are rarely valued in their own right. Rather, they are converted into something that enhances the subjective well-being of the person. Utility is thus a function of payoffs, and, importantly, such utility often accumulates in a multiplicative rather than additive manner.

B.8.2. *Additive Payoffs to Multiplicative Utility.* Imagine an interaction in which two people must divide up a resource to which they both have a claim (e.g., an inheritance, as in the main text). They coordinate using a particular norm of fairness and each receives her share of the resource as payoff. Assume this resource is physical, such as a cooking pot or a sum of money. It is unlikely that an individual will derive much utility from simply possessing a pot or a stack of paper money. Rather, she will use this payoff to enhance her well-being, e.g., by using it to obtain or process food that will feed herself and her family and thereby contribute to her and their survival. In this example, the utility received from a payoff is the resultant probability of survival of the receiver in the current time step. However, unlike the actual interaction payoff, survival (like measures of fitness in evolutionary models) is time dependent, and is thus a multiplicative rather than an additive function. The probability of surviving to some time t depends on the probability that you survived to time $t-1$, and therefore on the probability of survival in all previous time steps. For instance, if payoffs in each time step are constant and sufficient to ensure a probability of surviving from one time step to the next of 0.7, then the probability that a person survives until the third time step is $0.7 \cdot 0.7 \cdot 0.7 = 0.7^3 = 0.34$. Thus, utility is an exponential function of survival, and, consequently, an exponential function of payoff.

When utility is equated with the probability of survival, then it has a minimum of zero and maximum of one. However, there may be cases in which payoffs contribute to some subjective measure of utility that can take values up to $+\infty$. For instance, imagine that utility is directly proportional to a person's social standing, where social standing equals the number of others in the population who believe that the person is generous. Let a person's social standing increase when she offers support to others, and decrease to zero when she requests support. People may request support when their probability of survival to a given time step drops below 1, and the lower a person's probability of survival, the more likely she is to request support. When a person receives sufficient payoffs such that her probability of survival to time t is > 1, then she may offer her surplus payoffs to others as a form of support. Let a person's social standing decrease by a factor equal to the probability that she requests support from others in order to ensure her survival, and increase by a factor equal to the proportion of payoffs she earns in excess of those needed to ensure her own survival. Thus, if payoffs in each time step are constant and 30% more than are necessary to ensure survival, then, on average, a person's initial social standing is increased after three time steps by a factor of $(1+0.3)(1+0.3)(1+0.3) = 1.3^3 = 2.2$. Similarly, if payoffs are constant and only 70% of what is necessary to ensure survival, then, after three time steps, social standing will be, on average, 0.34 of its initial value (analogous to calculations for survival, above). Such exponential change of average social standing (utility) might occur if every individual starts life with social standing ≥ 1. A single request for support reduces social standing to 0, such that everyone who considers a person to be generous hears about the request and changes their mind permanently. The probability of requesting support is 1-prob(survival). If a request for support is made, social standing remains at zero indefinitely and cannot be recovered. In contrast, each time a person offers support, all of those individuals who already know of her generous reputation are inspired to spread this reputational information to one other naive individual with a probability of 1. The number of offers of support (or, if a fraction, the probability thereof) made by an individual at a given time is equal to her payoff in excess of 1.

B.8.3. *Modifying the Original Model.* In the model in the main text, it is assumed that individuals attempt to maximize their average perceived payoff. They do this, in a given updating step, by comparing the average perceived payoff of their current phenotype with those of other phenotypes to which they might change, and biasing their decision to change (or not) toward the phenotype with the highest average perceived payoff. The average perceived payoff $\tilde{w}(t)$, specific to a given phenotype, is the anticipated payoff to an individual with that phenotype at time t, averaged over the theoretically infinite ensemble of all individuals with that phenotype in that time step. This is the expectation (arithmetic mean payoff) for a given phenotype in a given time step, and is independent of any payoffs received in previous time steps. Thus, the model implicitly assumes that payoffs accumulate additively over the lifespan. Perceived payoff can take values ranging from $-m$ (e.g., $\tilde{w}_{S1X}$ when $a_S = c = p_{S22} = 1$) to $1 + b$ (e.g., $\tilde{w}_{S11}$ when $a_S = 0$ and $p_{L11} = 1$), in the absence of cultural group valuation ($i = 0$). Recall that $m \geq 0$ is the learning cost associated with cross-cultural competence, and is suffered by all cross-culturally competent phenotypes, and only them. Also recall that $b \geq 0$ is the extra payoff benefit associated with successfully coordinating with an out-group individual.

Using the example above, assume that, instead of attempting to maximize perceived payoffs, individuals attempt to maximize their social standing, which we can equate with utility (u). Let utility be a function of perceived payoff, such that

$$u(t) = S_0 \left[\tilde{w}(t) + m\right]^t \tag{A.54}$$

where S_0 is initial social standing at time $t = 0$, and $[\tilde{w}(t) + m] \in [0, +\infty]$ is the expected (arithmetic mean) proportional increase in social standing, i.e., the factor by which social standing at time t is multiplied. Note that adding the constant m standardizes the lower bound of this term at 0. For convenience, in what follows I will assume that initial social standing $S_0 = 1$.

If individuals are concerned only with maximizing utility in the current time step, they will attempt to maximize the arithmetic mean rate of increase in utility,

$$\hat{u}_a = \mathbb{E}\left[u\right] = \tilde{w} + m \tag{A.55}$$

which entails maximizing the perceived payoff $\tilde{w}$. If this is the case, then, in the original model, we can substitute $\hat{u}_a$ for $\tilde{w}$, for each respective phenotype in Table A.4 and Equations A.23 and A.24. Dynamics will be identical, as the only functional modification has been to add a constant (m) to all perceived payoffs. Differences in the arithmetic mean rate of utility increase are linear with respect to perceived payoff. For instance, the difference in $\hat{u}_a$ between a phenotype receiving $\tilde{w} = 1$ and another receiving $\tilde{w} = 2$ is the same as that between phenotypes receiving $\tilde{w} = 9$ and $\tilde{w} = 10$.

Now assume that, rather than attempting to maximize their utility within a given time step, individuals instead attempt to maximize their lifetime utility, i.e., their social standing across a hypothetically infinite number of time steps. This strategy takes advantage of the fact that utility is time dependent, i.e., utility in the current time step is a function of utility obtained in all previous time steps. For a given perceived payoff $\tilde{w}$, we can define utility as

$$u(t) = (\tilde{w} + m)^t = e^{\ln(\tilde{w}+m)\cdot t} \tag{A.56}$$

In order to maximize lifetime utility, one should attempt to maximize the rate of exponential growth of social standing, $\ln(\tilde{w} + m)$. Because this rate is constant for a given perceived payoff $\tilde{w}$, it is also the time-averaged (geometric mean) rate of increase in utility:

$$\hat{u}_g = \ln(\tilde{w} + m) \tag{A.57}$$

Note that, for any perceived payoff such that $(\tilde{w} + m) \neq 1$, expected utility (with growth rate $\hat{u}_a$) does not equal time-averaged utility (with exponential growth rate $\hat{u}_g$). Therefore, the utility function in Equation A.54 satisfies the definition of a non-ergodic function (Peters 2019).

As shown in Figure A.27, the rate of increase in geometric mean utility is not a linear function of payoff. When attempting to maximize $\hat{u}_g$, low payoffs are disproportionately undervalued, and increasing payoffs yield diminishing returns. Intuitively, this means that people are very averse to low payoffs, as they risk permanent loss of their social standing. In contrast, when payoffs are high, social standing accumulates exponentially over time. However, further increases to the rate of exponential growth provide only nominal gains to the accumulation of social standing over time. To incorporate this non-ergodic utility into the original model, substitute $\hat{u}_g$ for $\tilde{w}$, for each respective phenotype in Table A.4 and Equations A.23 and A.24.

B.8.4. *Model Analysis.* The dynamics of this model, as shown in Figure A.28, are broadly similar to those of the original model, with a few notable exceptions. In the situation of interest, such that members of a minority low-power group (S) engage in frequent inter-group interaction (low a_S) with members of a culturally-distinct majority (L), the top row of Figure A.28 shows that the cross-culturally competent phenotype $S2X$ will reach high frequency in group S when cultural group identity does not contribute to payoff ($i = 0$). As explained in the main text, if this occurs, S-typical norm 1 is likely to be lost from the population after one generation. However, if group identity

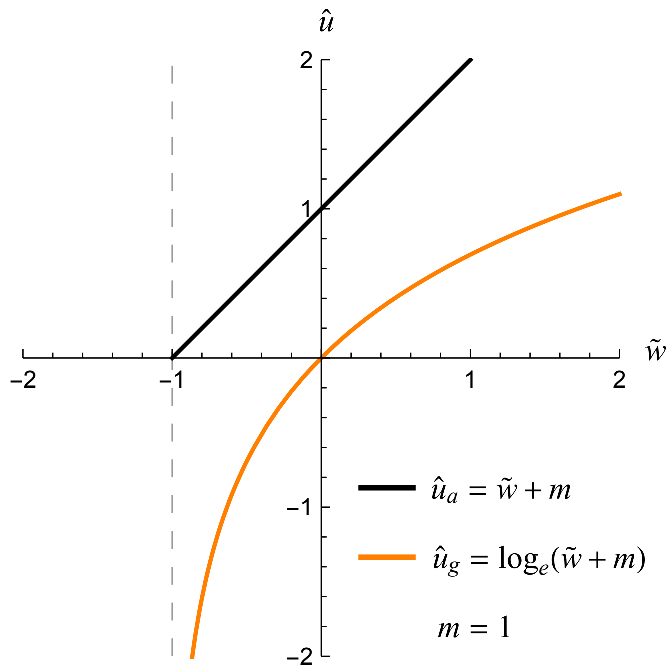

FIGURE A.27. Arithmetic (black) and geometric (orange) mean rate of increase in utility as a function of average perceived payoff $\tilde{w}$ and the constant m. The minimum possible value of $\tilde{w}$ is $-m$.

is valued ($i > 0$), the cross-cultural phenotype $S1X$ can attain high frequency (bottom row of Figure A.28), and norm 1 can potentially be preserved in group S indefinitely (see main text). Note that the uni-culturally competent phenotype $S11$ is quickly lost from group S. These dynamics are very similar to those in the original model, and demonstrate that the main qualitative results are robust to changes in the multiplicative versus additive nature of the utility function.

One major difference between this model with non-ergodic utility and the original is that it is considerably more difficult to maintain the cross-culturally competent $S1X$ phenotype in group S (e.g., it requires a higher value of i), especially when the payoff bias μ is low. This is because the non-ergotic utility function causes high payoffs to be undervalued. Therefore, if $S1X$ has a high perceived payoff that is greater than that of $S2X$, the difference between these utilities (payoffs) will be

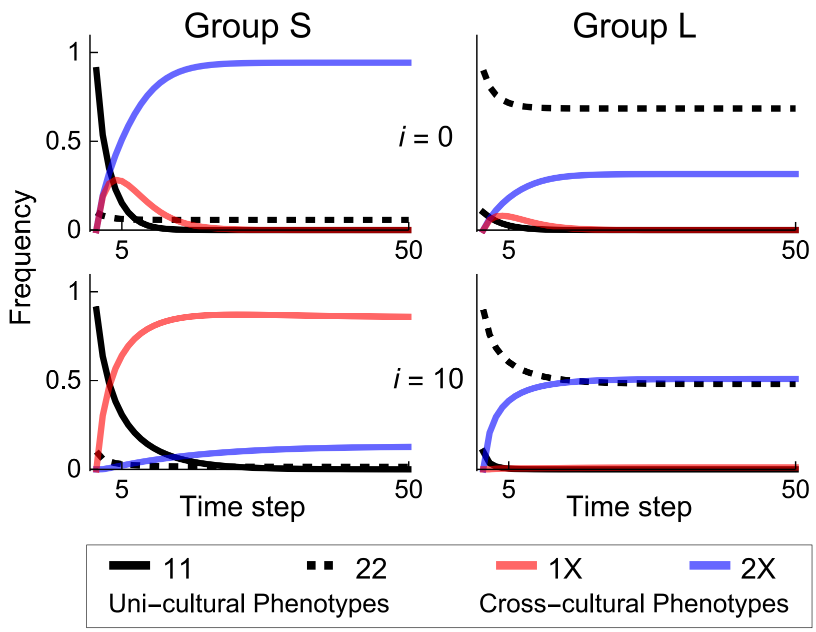

FIGURE A.28. Model simulations of phenotype frequency trajectories for the low-power minority group (S) and high-power majority group (L) simulated from the full model with a non-ergodic utility function where: $(b_S, b_L) = (1, 0)$, $(a_S, a_L) = (0.4, 0.7)$, $c = 0.1$, $\mu = 2$, $m = 1$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$, $p_{S22} = p_{L11} = 0.1$, given $i = 0$ (top row) and $i = 10$ (bottom row). Compare to Figure 2A.

reduced compared to the original model. When inter-group interaction is high (low a_S), $S1X$ often suffers the cost c, causing it to re-evaluate its phenotype in the updating phase. During these reevaluations, there is a higher probability (though still less than 0.5) that $S1X$ transitions to $S2X$, even if it has a higher perceived payoff. Because norm 2 occurs at high frequency in group L, $S2X$ reevaluates its phenotype less often than $S1X$, and so transitions to $S1X$ occur less often than $S1X \to S2X$.

A comparison of Figures A.29 and A.30 with Figures A.11 and A.12, reveals that the dynamics of the present model and the original model respond in similar ways to changes in the values of several model parameters, though there are differences worth noting. Figures A.29A-D and A.30A and B, show that this model responds in similar ways to changes in rates of inter-group interaction (a) and valuation of group identity (i).

In contrast to the original model, increasing the power difference between groups S and L (high b_S) is detrimental to the sustainability of $S1X$ (Figures A.29E and A.30C). This is because high b_S increases the payoffs to $S1X$ and $S2X$, thereby decreasing the utility difference between them and increasing the probability of transitions $S1X \rightarrow S2X$, as described above.

In contrast to the original model, increasing the cost of cognitive dissonance (c) is beneficial to the maintenance of $S1X$ (Figures A.29F and A.30D). This is because, initially, $S2X$ suffers c often while $S11$ is still at high frequency. High c slows the transition of $S11$ to $S1X$ and $S2X$, because it decreases the payoffs of these cross-culturally competent phenotypes relative to $S11$. Compared to the original model, $S1X$ suffers c less often because $L2X$ attains higher frequency in group L than in the original model. This is because the non-ergodic payoff function reduces the effect of the payoff advantage of $L22$ over $L2X$, resulting in a higher probability of transitions $L22 \rightarrow L2X$.

Similar to the original model, increasing the learning cost of cross-cultural competence (m) is slightly detrimental to the sustainability of $S1X$ when inter-group interaction and valuation of ethnic identity are high (Figures A.29G and A.30E). For reasons similar to the original model, high m decreases the rate of $S11 \rightarrow S1X$, but does not decrease the rate of $S1X \rightarrow S2X$. Also similar to the original model, maintenance of $S1X$ is enhanced by increasing the bias toward higher utility/payoffs (μ) during individuals' decisions to modify their phenotype (Figures A.29H and A.30F). This is because high μ counteracts the effect of the non-ergodic utility function, and makes individuals more sensitive to small differences in utility. Thus, initially, when $\tilde{w}_{A1X}$ is high and greater than $\tilde{w}_{A1X}$, yet the difference in their utilities is low, high μ will reduce the likelihood of transitions $S1X \rightarrow S2X$.

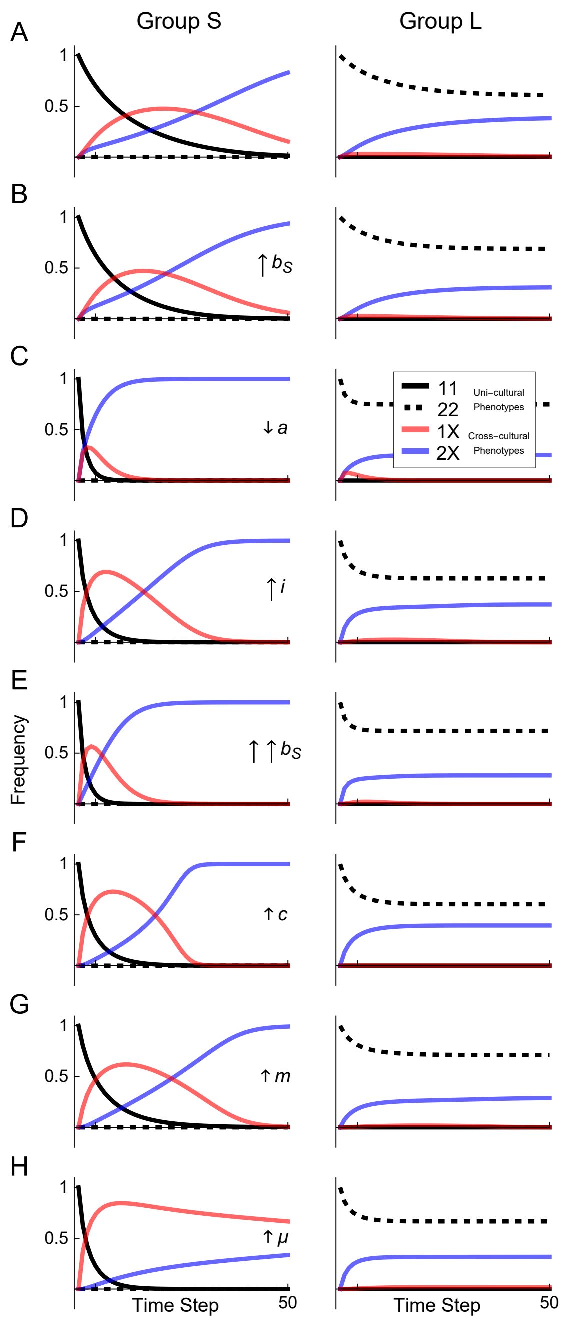

FIGURE A.29. (Caption next page.)

FIGURE A.29. Sensitivity of phenotype trajectories for the minority group (S) and majority group (L), simulating from the full model with a non-ergodic utility function. **A)** No power difference between groups, high in-group affinity: $(b_S, b_L) = (0, 0)$, $(a_S, a_L) = (0.8, 0.9)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, $i = 0$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$ and $p_{S22} = p_{L11} = 0.1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Low affinity: sames as **B**, except $(a_S, a_L) = (0.4, 0.7)$. **D)** Identity valuation: sames as **C**, except $i = 3$. **E)** Large power difference: same as **D**, except $(b_S, b_L) = (5, 1)$. **F)** High cognitive dissonance: sames as **D**, except $c = 0.9$. **G)** High cross-cultural learning cost: same as **D**, except $m = 3$. **H)** High payoff-bias for copying: sames as **D**, except $\mu = 3$. Compare to Figure A.11.

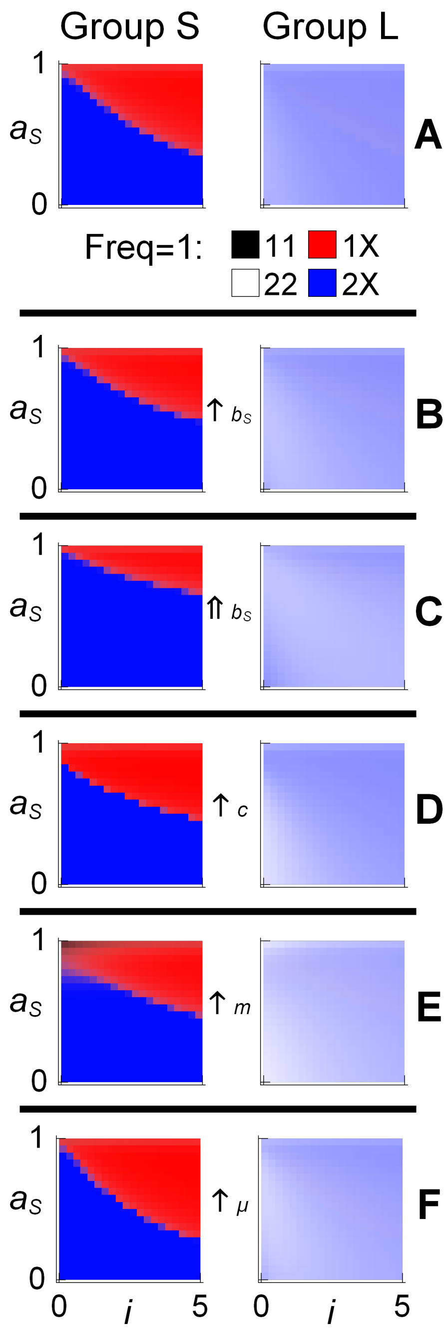

FIGURE A.30. (Caption next page.)

FIGURE A.30. Sensitivity of long-run phenotype frequencies for the minority group (S) and majority group (L), simulating from the full model with a non-ergodic utility function for 100 time steps. **A)** No power difference between groups: $(b_S, b_L) = (0, 0)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, and initial phenotype frequencies $p_{S11} = p_{L22} = 1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Large power difference: sames as **B**, except $(b_S, b_L) = (5, 1)$. **D)** High cognitive dissonance: sames as **B**, except $c = 0.9$. **E)** High cross-cultural learning cost: sames as **B**, except $m = 3$. **F)** High payoff-bias for copying: same as **B**, except $\mu = 3$. Compare to Figure A.12.

B.9. **Model with Stochastic Perception Error.**

B.9.1. *Model Design.* Here I modify the original full model by adding random error to people's average perceptions of the frequencies with which norms are used by both in- and out-group members ($\tilde{p}$'s). This has the effect of adding random error to people's perceptions of the average payoff that each phenotype will receive in the next time step ($\tilde{w}$'s). Such error therefore often affects people's norm adoption decisions in ways that are detrimental to their average payoff. Furthermore, by affecting people's norm adoption decisions, such error adds random noise to the frequencies of phenotypes in each time step. The question of interest is whether the main results of the original model analysis are robust to plausible levels of such error. For instance, can the cross-culturally competent phenotype $S1X$ be preserved in minority group S at equilibrium, given sufficient valuation of group identity (i)?

Stochastic error is introduced into Equations A.17 and A.18, which give, respectively, the frequency of a preference for norm 1 among group S members and the frequency of a preference for norm 2 among group L members, as inferred by a member of group S. These equations are modified as follows:

$$\begin{aligned}\tilde{p}_{S1in} = {} & p_{S11}(p_{S11} + p_{S1X} + p_{S2X} + p_{S22}) + \\ & p_{S1X}(p_{S11} + p_{S1X} + \tfrac{1}{2}p_{S2X}) + p_{S2X}(p_{S11} + \tfrac{1}{2}p_{S1X}) + \\ & N(0, \sigma^2)\end{aligned} \tag{A.58}$$

$$\begin{aligned}\tilde{p}_{L2out} = {} & p_{L22}(p_{S22} + p_{S2X} + p_{S1X} + p_{S11}) + \\ & p_{L2X}(p_{S22} + p_{S2X} + \tfrac{1}{2}p_{S1X}) + p_{L1X}(p_{S22} + \tfrac{1}{2}p_{S2X}) + \\ & N(0, \sigma^2)\end{aligned} \tag{A.59}$$

where $N(0, \sigma^2)$ represents a value drawn at each time step from a Normal distribution with mean 0 and variance σ^2, under the constraints $\tilde{p}_{S1in} \in [0, 1]$ and $\tilde{p}_{L2out} \in [0, 1]$. The probabilities $\tilde{p}_{S2out}$ and $\tilde{p}_{L2in}$ are found by reversing all group and norm indices in the subscripts of Equations A.59 and A.58, respectively.

B.9.2. *Model Analysis.* As shown in Figures A.31 and A.32, the dynamics of this model are broadly similar to those of the original model in the short term, responding in similar ways to changes in the values of model parameters (compare Figure A.32 to A.11). However, over the long term, it is much more difficult to sustain the cross-culturally competent phenotype $S1X$ in minority group S when there is stochastic error in people's perceptions about the frequencies of phenotypes in the in- and out-group (compare Figure A.31A and B to Figure 2A). This is a consequence of the fact that the probability of transitioning from one phenotype to another is modeled as a logistic function of the difference in anticipated payoffs between the current and potential phenotypes (see Table A.4). Thus, the probability of adopting the phenotype with the highest perceived average payoff is increased only slightly by stochastic error that increases the perceived difference in payoff between this phenotype and another (i.e., the rate of change in probability decreases as the perceived payoff difference increases). However, the probability of adopting the phenotype with the highest perceived average payoff is decreased relatively more by stochastic error of the same magnitude that decreases this perceived payoff difference (i.e., the rate of change in probability increases as the perceived payoff difference approaches 0).

For instance, in Figure A.31B when i is sufficiently large, $S1X$ initially has a higher average perceived payoff than $S2X$ in group S. However, whenever $S1X$ engages in inter-group interaction with $L22$ or $L2X$, there is a non-zero probability of a transition $S1X \rightarrow S2X$. This maintains $S2X$ at a non-zero frequency in group S (apparent in the lower row of Figure 2A). Stochastic perception error that increases the perceived average payoff difference between $S1X$ and $S2X$, decreases the probability of $S1X \rightarrow S2X$. However, stochastic error that decreases the perceived average payoff difference between $S1X$ and $S2X$, increases this transition probability to a relatively greater extent. Thus, stochastic perception error has the overall effect of increasing the probability of transitions $S1X \rightarrow S2X$. Over time, this increases the frequency of $S2X$ in group S, which eventually leads to a payoff advantage of $S2X$ over $S1X$, and the eventual extinction of $S1X$. This dynamic appears to be inevitable for the parameter conditions of interest. Figure A.31C and D show that the loss of $S1X$ can be slowed by increasing the valuation of group identity (i), which initially increases the perceived payoff advantage of $S1X$ over $S2X$, and by increasing the payoff bias in phenotype adoption (μ), which decreases the non-zero probability of $S1X \rightarrow S2X$ when $S1X$ has a higher perceived average payoff than $S2X$. However, even in such cases, the frequency of $S1X$ eventually degrades. Note that, stochastic error does not contribute to the loss of $S2X$ when it reaches high frequency in group S. This is because, when norm 2 reaches high frequency in both groups, $S2X$ individuals rarely receive less than the maximum expected payoff from any interaction, and therefore rarely consider changing their phenotype. Thus as the frequency of $S2X$ approaches 1, the probability of the transition $S2X \rightarrow S1X$ approaches 0.

Figure A.31E shows that, when stochastic perception error is very high, $S1X$ is quickly lost from the population. Figure A.31F shows that, in such a case, increasing i and μ have only negligible effects on this dynamic.

Another important implication of this model is the interaction between stochastic perception error and group identity valuation (i). As demonstrated in Figures A.31A and B and A.32C and D, i amplifies the strength of perception error. The reason for this can be seen in the expressions for perceived average payoffs (Equations A.19-A.22), where i multiplies the perceived frequency of a norm within the in-group and the perceived rarity of a norm within the out-group.

Thus, if our objective is to preserve a distinctive minority cultural norm (norm 1) through the sustainability of the cross-culturally competent phenotype $S1X$, this model suggests that a strategy relying on increasing the valuation of group identity must also attempt to minimize erroneous perceptions that norm 1 is less common than it is in reality. Such misconceptions contribute to the loss of $S1X$, even when it occurs at very high frequency.

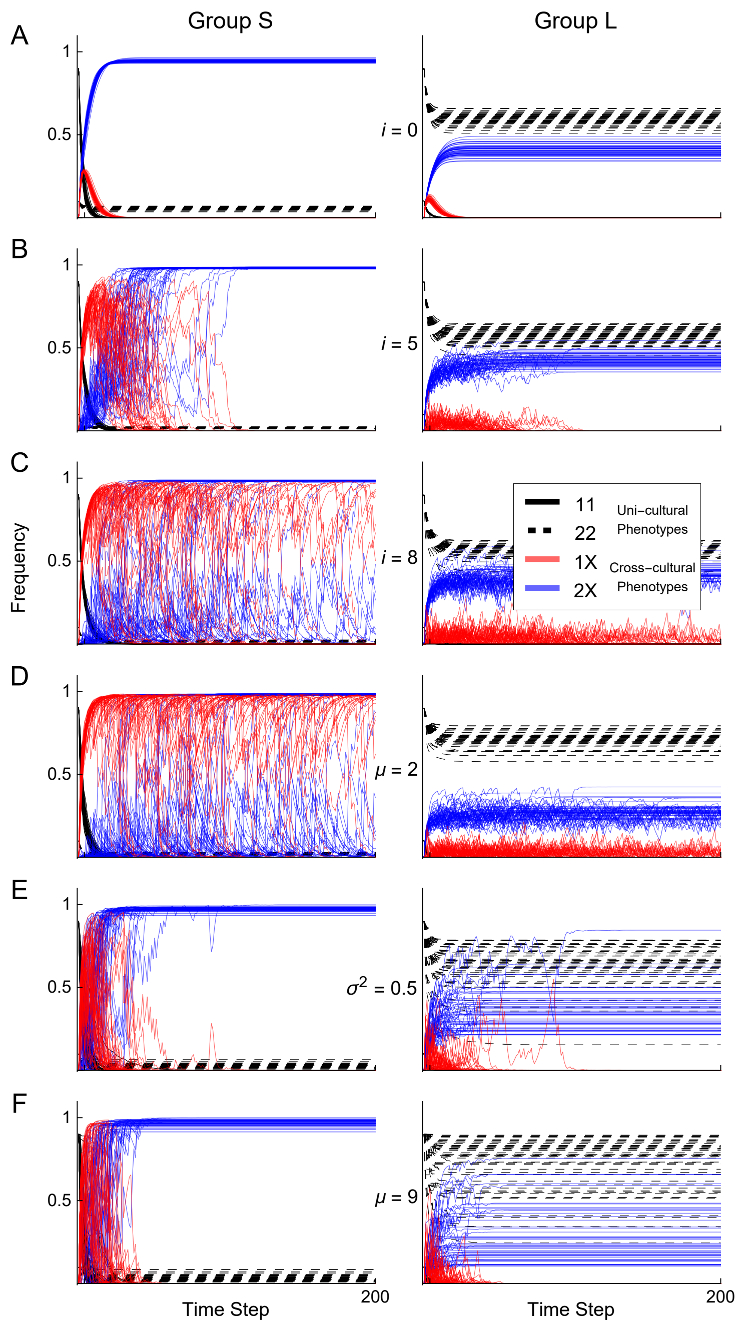

FIGURE A.31. (Caption next page.)

FIGURE A.31. Sensitivity of phenotype trajectories for the minority group (S) and majority group (L), simulating from the full model with stochastic perception error. In each plot, 50 simulations of the four phenotype trajectories are shown. Note that these simulations are run for 200 time steps, rather than 50 as in previous plots. **A)** Parameter conditions identical to the top row of Figure 2A, with $(b_S, b_L) = (1, 0)$, $(a_S, a_L) = (0.4, 0.7)$, $c = 0.1$, $\mu = m = 1$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$, $p_{S22} = p_{L11} = 0.1$, given $i = 0$. Here stochastic error variance is moderate: $\sigma^2 = 0.25$ (see Equations A.58 and A.59). **B)** Parameter conditions identical to **A**, except $i = 5$. Compare to the bottom row of Figure 2A. **C)** Higher group identity valuation: same as **B**, except $i = 8$. **D)** Higher payoff bias: same as **B**, except $\mu = 2$. **E)** Larger stochastic error: same as **D**, except $\sigma^2 = 0.5$. **F)** Unsuccessful attempt to sustain cross-culturally-competent phenotype $S1X$ in group S: same as **E**, except $i = 10$ and $\mu = 9$.

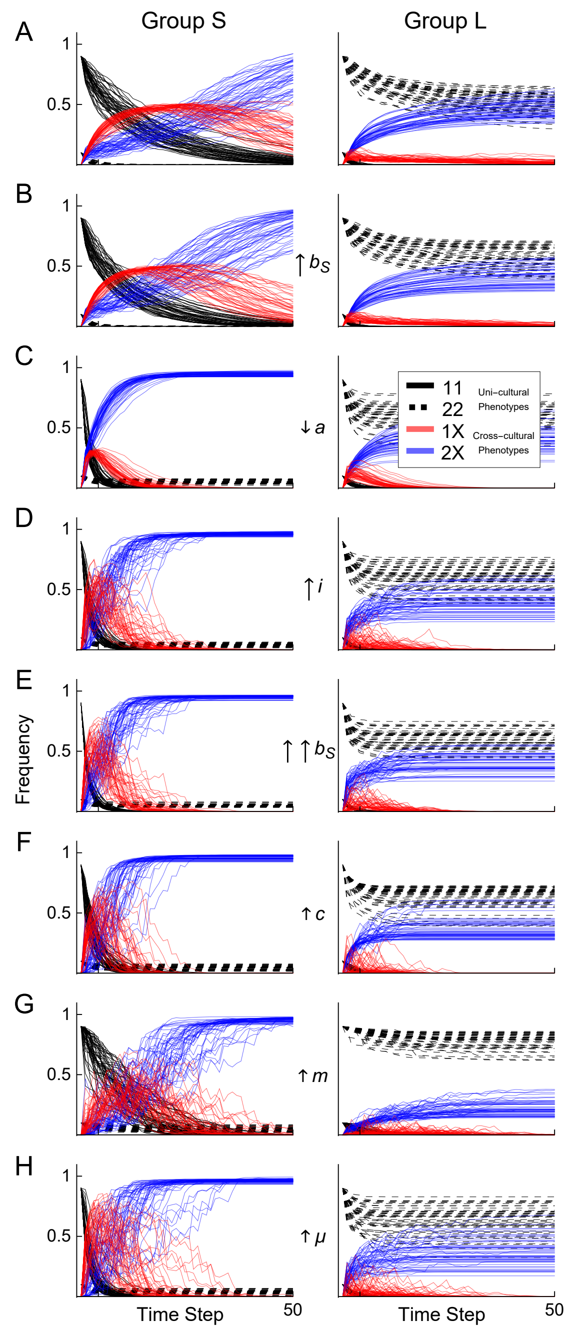

FIGURE A.32. (Caption next page.)

FIGURE A.32. Sensitivity of phenotype trajectories for the minority group (S) and majority group (L), simulating from the full model with stochastic perception error. In each plot, 50 simulations of the four phenotype trajectories are shown, all with $\sigma^2 = 0.5$ (see Equations A.58 and A.59). **A)** No power difference between groups, high in-group affinity: $(b_S, b_L) = (0, 0)$, $(a_S, a_L) = (0.8, 0.9)$, $c = 0.1$, $\mu = 2$, $m = 0.5$, $i = 0$, and initial phenotype frequencies $p_{S11} = p_{L22} = 0.9$ and $p_{S22} = p_{L11} = 0.1$. **B)** Power difference: same as **A**, except $(b_S, b_L) = (1, 0)$. **C)** Low affinity: sames as **B**, except $(a_S, a_L) = (0.4, 0.7)$. **D)** Identity valuation: sames as **C**, except $i = 1$. **E)** Large power difference: same as **D**, except $(b_S, b_L) = (5, 1)$. **F)** High cognitive dissonance: sames as **D**, except $c = 0.9$. **G)** High cross-cultural learning cost: same as **D**, except $m = 3$. **H)** High payoff-bias for copying: sames as **D**, except $\mu = 3$. Compare to Figure A.11.

References

Advani, A., Reich, B., 2015. Melting pot or salad bowl: The formation of heterogeneous communities. IFS Working Papers W15/30. London. URL: http://hdl.handle.net/10419/145442.

Akerlof, G.A., Kranton, R.E., 2000. Economics and Identity. The Quarterly Journal of Economics 115, 715–753.

Baer, G., 2004. ¿Comes carne de armadillo? acerca del problema de la identidad de los Matsiguenka del Bajo Urubamba en el Oriente Peruano. Boletín de Antropología (Universidad de Antioquia) 18, 317–336.

Bicchieri, C., 2006. The Grammar of Society: The Nature and Dynamics of Social Norms. Cambridge University Press, Cambridge.

Bisin, A., Patacchini, E., Verdier, T., Zenou, Y., 2011. Formation and persistence of oppositional identities. European Economic Review 55, 1046 – 1071.

Bisin, A., Verdier, T., 2001. The economics of cultural transmission and the dynamics of preferences. Journal of Economic Theory 97, 298 – 319.

Boyd, R., Richerson, P.J., 1985. Culture and the Evolutionary Process. The University of Chicago Press, Chicago.

Boyd, R., Richerson, P.J., 2009. Voting with your feet: Payoff biased migration and the evolution of group beneficial behavior. Journal of

Theoretical Biology 257, 331–339.

Bruner, J.P., 2019. Minority (dis)advantage in population games. Synthese 196, 413–427.

Bunce, J.A., 2020. Field evidence for two paths to cross-cultural competence: implications for cultural dynamics. Evolutionary Human Sciences 2, 1–16.

Bunce, J.A., McElreath, R., 2017. Interethnic interaction, strategic bargaining power, and the dynamics of cultural norms: A field study in an Amazonian population. Human Nature 28, 434–456.

Bunce, J.A., McElreath, R., 2018. Sustainability of minority culture when inter-ethnic interaction is profitable. Nature Human Behaviour 2, 205–212.

Carvalho, J.P., 2017. Coordination and culture. Economic Theory 64, 449–475.

Cohen, D., Nisbett, R.E., Bowdle, B.F., Schwarz, N., 1996. Insult, aggression, and the southern culture of honor: An "experimental ethnography.". Journal of Personality and Social Psychology 70, 945–960.

Erten, E.Y., van den Berg, P., Weissing, F.J., 2018. Acculturation orientations affect the evolution of a multicultural society. Nature Communications 9, 58.

Fantini, A.E., 2009. Assessing intercultural competence: Issues and tools, in: Deardorff, D.K. (Ed.), The SAGE Handbook of Intercultural Competence. SAGE, Los Angeles. book section 27, pp. 456–476.

Festinger, L., 1962. A Theory of Cognitive Dissonance. Stanford University Press, Stanford.

Gans, H.J., 1979. Symbolic ethnicity: the future of ethnic groups and cultures in America. Ethnic and Racial Studies 2, 1–20.

Gluckman, M., 1963. Papers in honor of Melville J. Herskovits: Gossip and scandal. Current Anthropology 4, 307–316.

Gow, P., 1993. Gringos and wild indians: images of history in western amazonian cultures. L'Homme 33, 327–347.

Gurven, M., Zanolini, A., Schniter, E., 2008. Culture sometimes matters: Intra-cultural variation in pro-social behavior among Tsimane Amerindians. Journal of Economic Behavior & Organization 67, 587–607.

Henrich, J., Boyd, R., 2008. Division of labor, economic specialization, and the evolution of social stratification. Current Anthropology 49, 715–724.

Hoffmann, R., 2006. The cognitive origins of social stratification. Computational Economics 28, 233–249.

Kandler, A., Unger, R., Steele, J., 2010. Language shift, bilingualism and the future of Britain's Celtic languages. Philosophical Transactions of the Royal Society of London B: Biological Sciences 365, 3855–3864.

Katz, D., Allport, F., 1931. Students' Attitudes: A Report of the Syracuse University Reaction Study. Craftsman Press.

Kuran, T., 1995. Private Truths, Public Lies: The Social Consequences of Preference Falisification. Harvard University Press, Cambridge.

Kuran, T., Sandholm, W.H., 2008. Cultural integration and its discontents. The Review of Economic Studies 75, 201–228.

Llosa Isenrich, E., Nieto Degregori, L., 2003. El Manu a Través de la Historia. PRO-MANU, Lima.

McElreath, R., 2016. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Texts in Statistical Science, CRC Press, Boca Raton.

McElreath, R., 2020. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. Texts in Statistical Science. second ed., CRC Press, Boca Raton.

McElreath, R., Boyd, R., 2007. Mathematical Models of Social Evolution: A Guide for the Perplexed. University of Chicago Press, Chicago.

McElreath, R., Boyd, R., Richerson, P.J., 2003. Shared norms and the evolution of ethnic markers. Current Anthropology 44, 122–129.

Medin, D., Ross, N., Cox, D., Atran, S., 2007. Why folkbiology matters: Resource conflict despite shared goals and knowledge. Human Ecology 35, 315–329.

Mesoudi, A., 2018. Migration, acculturation, and the maintenance of between-group cultural variation. PLOS ONE 13, 1–23.

Mohseni, A., O'Connor, C., Rubin, H., 2019. On the emergence of minority disadvantage: testing the cultural Red King hypothesis. Synthese doi:10.1007/s11229-019-02424-1.

O'Connor, C., 2019. The Origins of Unfairness: Social Categories and Cultural Evolution. Oxford University Press, Oxford, UK.

Olcina, G., Panebianco, F., Zenou, Y., 2018. Conformism, Social Norms and the Dynamics of Assimilation. IZA Discussion Papers 11436. Bonn. URL: http://hdl.handle.net/10419/180454.

Peters, O., 2019. The ergodicity problem in economics. Nature Physics 15, 1216–1221.

Portes, A., Rumbaut, R.G., 2014. Immigrant America: A Portrait. 4 ed., University of California Press.

Prentice, D.A., Miller, D.T., 1993. Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm.

Journal of Personality and Social Psychology 64, 243–256.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Rubin, H., O'Connor, C., 2018. Discrimination and collaboration in science. Philosophy of Science 85, 380–402.

Shepard, G.H., 2002. Three days for weeping: dreams, emotions, and death in the Peruvian Amazon. Medical Anthropology Quarterly 16, 200–229.

Shepard, G.H., Rummenhoeller, K., Ohl, J., Yu, D.W., 2010. Trouble in paradise: Indigenous populations, anthropological policies, and biodiversity conservation in Manu National Park, Peru. Journal of Sustainable Forestry 29, 252–301.

Spitzberg, B.H., Changnon, G., 2009. Conceptualizing intercultural competence, in: Deardorff, D.K. (Ed.), The SAGE Handbook of Intercultural Competence. SAGE, Los Angeles. book section 1, pp. 2–52.

Stan Development Team, 2017. Stan Modeling Language: User's Guide and Reference Manual, Version 2.17.0. Stan Development Team.

Stan Development Team, 2018. RStan: the R interface to Stan. URL: http://mc-stan.org/.

Wiessner, P., 2005. Norm enforcement among the Ju/'hoansi Bushmen. Human Nature 16, 115–145.