

**THE CULTURAL BRAIN HYPOTHESIS
AND THE TRANSMISSION AND EVOLUTION OF CULTURE**

by

Michael Muthukrishna

B.A., B.Eng. (Hon I), The University of Queensland, 2010

M.A., The University of British Columbia, 2012

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Psychology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2015

© Michael Muthukrishna, 2015

Abstract

Humans are an undeniably remarkable species with massive brains, amazing technology, and large, well-connected social networks. The co-occurrence of these traits is no accident. Here, I introduce the Cultural Brain Hypothesis and Cumulative Cultural Brain Hypothesis. The Cultural Brain Hypothesis is a single theory that explains the increase in brain size across many taxa. In doing so, it makes predictions about the relationships between brain size, adaptive knowledge, group size, social learning, and the length of the juvenile period. These predictions are consistent with existing empirical literature, tying together these otherwise disparate measured relationships. The Cumulative Cultural Brain Hypothesis makes predictions about the conditions under which these processes lead to a positive feedback loop between brain size and adaptive knowledge, ratchetting both upward in a co-evolutionary duet. I argue that these conditions, which include a rich environment, low reproductive skew, and high transmission fidelity, are the key to the uniquely human pathway, explaining our large brains and various aspects of our psychology that led to and sustain those large brains. The predictions of these two hypotheses are consistent with other theories within a Dual Inheritance Theory framework – the idea that natural selection led our species to develop a line of cultural inheritance in addition to the line of genetic inheritance shared by all species on Earth. I test many of these theories against competing theories across 4 experiments with human subjects. These experiments include cultural transmission experiments, which support a causal relationship between sociality and cultural complexity, and two social learning experiments, which test the environmental and individual predictors of biased social learning. The overall findings support Dual Inheritance Theory and the Cumulative Cultural Brain Hypothesis. Finally, I lay the groundwork for improving the way in which these theories account for human social structures by proposing a theory to help explain the social structures unique to our species. I then demonstrate the utility of this model by

using it to make predictions about the implications of population-level differences in personality and social influence for the transmission and evolution of culture.

Preface

The various models and experiments reported here are collaborative projects. In each, however, I am the primary author and am primarily responsible for the ideas, design, analyses, and authorship.

Chapters 1 and 6, the Introduction and Conclusion, are reviews and commentary on my research area and its connection to the broader field. I am the sole author in both cases.

Chapter 2 is a manuscript currently in preparation. Author order is Muthukrishna, M., Chudek, M., and Henrich, J. The research is a theoretical model and therefore does not require ethical approval.

Chapter 3 is based on a paper published in the journal *Proceedings of the Royal Society B: Biological Sciences*. The official citation that should be used when referencing this material is: Muthukrishna, M., Shulman, B. W., Vasilescu, V., & Henrich, J. (2013). Sociality influences cultural complexity. *Proceedings of the Royal Society B: Biological Sciences*, 281(1774). The research in this chapter was approved by the Behavioural Research Ethics Board of UBC (BREB certificate #H10-03272).

Chapter 4 is based on a paper forthcoming in the journal *Evolution and Human Behavior*. The official citation that should be used when referencing this material is: Muthukrishna, M., Morgan, T. J. H., & Henrich, J. (forthcoming). The when and who of social learning and conformist transmission. *Evolution and Human Behavior*. The research in this chapter was approved by the Behavioural Research Ethics Board of UBC (BREB certificate #H13-02185).

Chapter 5 is based on a paper currently under review. Author order is Muthukrishna, M. and Schaller, M. The research is a theoretical model and therefore does not require ethical approval.

Table of Contents

Abstract.....	ii
Preface	iv
Table of Contents	v
List of Tables	xi
List of Figures.....	xv
Acknowledgements	xxii
Dedication	xxiv
Chapter 1: Introduction	1
1.1 Culture Makes Us Human.....	3
1.1.1 Dual Inheritance Theory / Gene-Culture Coevolution & Cultural Evolution.....	4
1.2 Contributions of Present Research	5
Chapter 2: Cultural Brain Hypothesis & Cumulative Cultural Brain Hypothesis	7
2.1 Model Description.....	11
2.1.1 The Lifecycle.....	13
2.1.2 Stage 1: The Birth Stage	15
2.1.2.1 Adaptive Knowledge and the Number of Children.....	15
2.1.2.2 Genetic Transmission and Mutation	17
2.1.3 Stage 2: Learning	17
2.1.3.1 Stage 2A: Vertical Transmission or Asocial Learning.....	18
2.1.3.2 Stage 2B: Asocial learning, Oblique learning, or No update.....	18
2.1.4 Stage 3: Migration.....	19
2.1.5 Stage 4: Selection Based on Brain Size and Adaptive Knowledge.....	19

2.1.6	Summary.....	21
2.2	Theory.....	22
2.3	Cultural Brain Hypothesis.....	23
2.3.1	Predictions.....	24
2.3.2	Testing Predictions.....	26
2.3.2.1	Brain Size and Group Size.....	27
2.3.2.2	Brain Size and Social Learning.....	29
2.3.2.3	Brain Size and Juvenile Period.....	30
2.3.2.4	Group Size and Juvenile Period.....	32
2.3.2.5	Group Size and Adaptive Knowledge.....	34
2.3.2.6	Summary.....	35
2.3.3	Processes Underlying Empirical Relationships.....	35
2.3.3.1	Richness of the Ecology (λ).....	37
2.3.3.2	Reproductive Skew or Mating Structure (ϕ).....	41
2.3.3.3	Transmission Fidelity and Asocial Learning Efficacy (τ and ζ).....	44
2.4	Cumulative Cultural Brain Hypothesis.....	46
2.4.1	Transmission Fidelity Drives Larger Brains.....	48
2.4.2	Social Learning Benefits from Smart Ancestors.....	49
2.4.3	Mating Structure Matters.....	50
2.4.4	Asocial Learning Can Lead to Larger Brains.....	51
2.4.5	The Richness of the Environment Can Give You a Bigger Brain.....	52
2.5	General Discussion.....	52
2.5.1	Summary of Key Findings.....	52

2.5.2	The Model in the Context of Other Theory and Findings	53
2.5.2.1	Cultural Intelligence Hypothesis	56
2.5.3	Limitations and Future Directions.....	56
Chapter 3: Sociality Influences Cultural Complexity		58
3.1	Methods	59
3.2	Results	63
3.2.1	Experiment 1	63
3.2.2	Experiment 2	68
3.3	Discussion.....	71
Chapter 4: When and Who of Social Learning		75
4.1	Theoretical Research	76
4.2	Experimental Research	78
4.3	Present Research.....	78
4.4	Methods	80
4.4.1	Participants.....	80
4.4.2	General Design.....	81
4.4.3	Experiment 1: Number of Options.....	82
4.4.4	Experiment 2: Transmission Fidelity and Payoffs.....	83
4.4.5	Background Measures.....	84
4.5	Analysis	85
4.6	Results	88
4.6.1	Number of Options (Experiment 1).....	88
4.6.1.1	Social Learning.....	89

4.6.1.2	Conformist Bias	91
4.6.2	Transmission Fidelity and Payoffs (Experiment 2)	93
4.6.2.1	Social Learning	93
4.6.2.2	Conformist Bias	95
4.6.3	Individual Variation in Social Learning Strategies	98
4.6.3.1	Social Learning	98
4.6.3.2	Conformist Bias	100
4.7	Discussion	103
Chapter 5: Cultural Dispositions, Social Networks, and the Dynamics of Social Influence		109
5.1	Cultural Differences in Extraversion and Conformity	110
5.1.1	Implications for the Structure of Social Networks	112
5.1.2	Implications for Societal Outcomes of Social Influence Processes	114
5.1.2.1	Consolidation of Existing Opinion Majorities	115
5.1.2.2	Diffusion and Spread of New Ideas	115
5.1.2.3	Obvious and Non-obvious Effects of Cultural Differences	116
5.2	Computational Modeling of Social Interaction and Social Influence	117
5.2.1	Overview of Our Computational Modeling Methods	119
5.2.2	Simulation of Individual Differences and Cultural Differences	120
5.2.3	Phase 1: Emergent Differences in the Structure of Social Networks	124
5.2.4	Phase 2: Interpersonal Influence within Social Networks	129
5.3	Simulated Effects of Cultural Differences on Consolidation	131
5.3.1	Summary and Discussion	135
5.4	Simulated Effects of Cultural Differences on the Diffusion of Innovations	137

5.4.1	The "Lone Ideologue" Context.....	138
5.4.2	The "Ideologue Accompanied by Disciples" Context.....	141
5.4.3	Summary and Discussion.....	146
5.5	General Discussion.....	147
5.5.1	Implications for Real-World Populations.....	149
5.5.2	Empirical Testability of the Hypotheses.....	151
5.5.3	Novel Features of Modeling Methods Employed Here.....	152
5.5.4	Lacunae, Limitations, and Directions for Future Research.....	154
5.5.5	Broader Applications of These Modeling Methods.....	158
5.5.6	Envoi.....	159
Chapter 6: Conclusion		161
6.1	A Theory of Human Behavior.....	162
6.1.1	Building Theory.....	166
6.1.2	Testing Theory.....	168
6.1.3	Present Research Enterprise.....	169
6.2	Central Questions and Answers.....	173
6.2.1	Why Now?.....	173
6.2.2	Why Us?.....	175
6.2.3	What Psychology Do We Need For Culture?.....	176
6.2.4	What Sociality Do We Need For Culture?.....	178
6.2.5	How Does This Connect With Our Broader Psychology?.....	179
6.2.6	What Are The Other Central Questions?.....	180
6.3	Technology Shapes Our Theories.....	183

References.....	187
Appendices	223
Appendix A Supplementary Materials for Sociality Influences Cultural Complexity	223
A.1 Participants.....	223
A.2 Experimental Design	225
A.3 Further Results and Analyses	229
A.4 Normalized Cross-correlation Metric.....	235
A.5 Rater Training and Testing	236
A.6 Rating Scales.....	236
Appendix B Supplementary Materials for When and Who of Social Learning	241
B.1 Participants.....	241
B.2 Experimental Design	242
B.3 Background Measures.....	246
B.4 Further Analyses and Results	251

List of Tables

Table 2.1. Correlations between log mean brain size, log mean adaptive knowledge, log mean group size, mean social learning, and mean juvenile period across the entire parameter space. The table has been color coded from white ($r = 0$) to blue ($r = 1$) for ease of comprehension.26

Table 2.2. Correlations between log mean brain size, log mean adaptive knowledge, log mean group size, mean social learning, and mean juvenile period. Primarily asocial learners ($s < .5$) are in the bottom triangle and primarily social learners $s > .5$ are in the top triangle. The table has been color coded from red ($r = -1$) to blue ($r = 1$) for ease of comprehension.....26

Table 3.1 OLS regression of standardized image rating scores on the main effects and interaction of Generation and Treatment (1-Model/5-Model), controlling for Male (gender, male = 1) and Age (standardized). By alternating the dummy coding of treatment, we directly compare the effect of Generation by looking at the Generation coefficients. In the 5-Model treatment, image ratings improve by 0.23 standard deviations per generation. In contrast, in the 1 model treatment there is no significant improvement in image ratings (and a possible decline).....66

Table 3.2 Binary logistic regression of the presence or absence of each component of the target image in each participant's attempted image on the corresponding component in each of the 5 potential models. We control for non-independence between participant's image components using clustered robust standard errors. The odds ratios reported reveal a large and significant bias for the best model, but also biases for the 3 next best models. We control for Generation, Male and Age. .67

Table 3.3 OLS regression of standardized knot rating scores on the Generation and Treatment (1-Model/5-Model), and their interaction, controlling for Male, Age (standardized) and knot-tying experience. By alternating the dummy coding of treatment, Table 3.3 directly compare the effect of Generation by looking at the Generation coefficients. The loss of skill within both the first

3 generations and the last 7 generations is twice as fast in the 1-Model treatment compared to the 5-Model treatment. We conducted a test of joint significance of treatment and treatment-generation interaction by statistically comparing regression models with and without these variables. Results indicate a statistically significant effect of treatment and treatment-generation interaction.....70

Table 4.1 Binary logistic multilevel model of decision to switch regressed on the proportion of participants in the option (in 10% increments for easier interpretation), the reciprocal and number of options (separate models), and the number of participants in the group. All coefficients are odds ratios. We control for common variance created by multiple observations from the same person with random effects for each individual.90

Table 4.2 Binary logistic multilevel model of decision to switch to majority on majority size, transmission fidelity, payoff, and number of participants in the group. All coefficients are odds ratios. We control for common variance created by multiple observations from the same person with random effects for each individual.95

Table 4.3 OLS regression model percentage of decisions that were changed after viewing social information regressed on theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the reference group. These results show a negative relationship between IQ and social learning with higher IQ resulting in less social learning. The regression models reported show all theoretically inspired predictors; the regression model is significant when the non-significant predictors are removed (see Appendix B.4).99

Table 4.4 OLS regression model of standardized log measures of strength of conformist transmission (α) regressed on our theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the

reference group. These results suggest a consistent quadratic (U shaped) relationship between IQ and the strength of the conformist transmission bias. Both those who scored high or very low on the IQ test were more likely to have stronger conformist transmission biases than those who scored in the middle. In Experiment 1, which is arguably more sensitive than Experiment 2 because there are often more than 2 options, conformist biases strengthen among older individuals and weakens among males.102

Table 5.1 Structural properties of the social networks that emerged in Phase 1 of the simulations, as a function of the population-wide mean level of extraversion within the population. Tabled values are means computed across 100 simulations for each of the three levels of extraversion (standard deviations around these means are in parentheses).128

Table 5.2 Results of multiple regression analysis with random effects for each network, testing the effects that population-wide mean levels of extraversion and conformity had on the log of the number of influence opportunities that elapsed before majority opinion reached a 2/3 super-majority threshold. We used the log of number of influence opportunities due to positive skew in the residuals. **R²** calculated on full model, including random effects (Nakagawa & Schielzeth, 2013).135

Table 5.3 Results of binary logistic regression analysis with random effects for each network, testing the effects that population-wide mean levels of extraversion and conformity had on the likelihood that a new belief—held initially by just one highly extraverted “lone ideologue”— spread to 50% of the entire population. Pseudo-**R²** calculated on full model, including random effects (Nakagawa & Schielzeth, 2013).141

Table 5.4 Results (odds ratios) of binary logistic regression analyses with random effects for each network, testing the effects that population-wide mean levels of extraversion and conformity

had on the likelihood that a new belief— held initially by an ideologue along with disciples——
spread to 50% of the entire population. Each row presents results associated with the subset of 9000
simulations associated with a specific number of disciples (varying from 1 to 12).145

List of Figures

Figure 2.1. Simulation variables, parameters, and flowchart. Four life stages for all individuals.	14
Figure 2.2. Fitness cost as measured by death rate plotted for different brain sizes and adaptive knowledge. Red indicates a higher death rate.	21
Figure 2.3. Histogram of mean social learning probability (s) across all simulations. Note that the relative count size of the two regimes is a reflection of the range of parameters we chose rather than a reflection of the world. We chose parameters that encompassed a realistic range, but allowed us to investigate the threshold between asocial and social learning (e.g. transmission fidelity values greater than 75% rather than 0% to 100%).	25
Figure 2.4. Example of correlations between brain size and group size for regimes that emerge in our model (left). Correlation between brain size and group size in the empirical literature (right; Barton, 1996).	28
Figure 2.5. Example of correlations between brain size and incidences of social learning for regimes that emerge in our model (left). Correlation between brain size and incidences of social learning in the empirical literature (right; Reader & Laland, 2002).....	30
Figure 2.6. Example of correlations between brain size and the length of the juvenile period for regimes that emerge in our model (left). Correlation between brain size and juvenile period in the empirical literature (right; Joffe, 1997).	32
Figure 2.7. Example of correlations between group size and the length of the juvenile period for regimes that emerge in our model (left). Correlation between mean group size and relative juvenile period in the empirical literature (right; Joffe, 1997).....	33

Figure 2.8. Example of correlations between group size and the mean adaptive knowledge for regimes that emerge in our model (left). Female group size and number of cultural traits in the empirical literature (right; Lind & Lindenfors, 2010).35

Figure 2.9. Causal relationships suggested by Cultural Brain Hypothesis.....36

Figure 2.10 Bean plots showing the distribution of (a) brain size and (b) social learning means for different values of λ . The dotted horizontal line shows the global mean and the bolded horizontal lines show the group means. Bean plots show the distribution of values. We see 3 regimes emerge for brain size. The regime with the largest brain size corresponds to the realm of cumulative cultural evolution (see Section 2.4). (c) Plot showing the rate of extinction for different values of λ . Note that in all cases, $\lambda = 2$ has not yet reached equilibrium.39

Figure 2.11 (a) Mean brain size showing the encephalization slope for different values of λ . Higher values have a larger encephalization slope. Note that brain size initially shrinks before growing again. This shrinkage corresponds to the transition from asocial to social learning, resulting in more efficient brains – that is smaller brains are needed to acquire more adaptive knowledge if that knowledge is acquired socially. There is some evidence that the human brain has been shrinking in the last 10,000 to 20,000 years, which may be evidence that our species is not at equilibrium (Henneberg, 1988). (b) Mean social learning over generations. There are two features of note. First, we started our simulations with social learning to show that social learning is maladaptive in the absence of adaptive knowledge. Asocial learners quickly invade. It is only when asocial learners have generated sufficient adaptive knowledge that social learners again have an advantage. Since we know that two regimes reliably emerge, mean social learning in these plots represents the relative number of conditions in which social and asocial learners emerge rather than a value of social learning characteristic of the world.40

Figure 2.12 Bean plots showing the distribution of (a) brain size and (b) social learning means for different values of ϕ . The dotted horizontal line shows the global mean and the bolded horizontal lines show the group means. Bean plots show the distribution of values. W (c) Plot showing the rate of extinction for different values of ϕ43

Figure 2.13 Bean plots showing the distribution of social learning for different values of τ and ζ . The dotted horizontal line shows the global mean and the bolded horizontal lines show the group means. Bean plots show the distribution of values. Unsurprisingly, higher transmission fidelity leads to more social learning. However, we see an interaction, where social learning is more likely to evolve when asocial learning efficacy is higher. These simulations exaggerate this overall pattern, because we begin with social learners, but the key message is that social learners stand on the shoulders of effective asocial learners whose knowledge they exploit.....45

Figure 2.14 Log mean brain size against the probability of acquiring the mean adaptive knowledge in the group via asocial learning for $\lambda = 1$. Circle size indicates the mean population size. Red simulations did not cross the threshold into cumulative cultural evolution, but aquamarine simulations did. In the lower graph, we zoom into the 0% to 1% probability range.47

Figure 2.15 Percentage of simulation which enter the realm of cumulative cultural evolution.48

Figure 2.16 Percentage of simulations which enter the realm of cumulative cultural evolution.....51

Figure 3.1 (a) An illustration of the experimental design. (b) The target image for Experiment 1. Note the words “Forty Two” at the base of the image and the red glow around these words and the circle. Participants were not required to recreate the dimension arrows. (c) The knots

used in Experiment 2. Participants were asked to tie this setup to two chairs. Larger versions of (b) and (c) can be found in Appendix A, Figures A.1 and A.2.60

Figure 3.2 Mean Image editing skills over 10 generations for the 1-Model and 5-Model treatments in Experiment 1. Scores rescaled between 0 and 100, where 100 is a perfect score. Linear lines of best fit emphasize a cumulative improvement in the 5-Model treatment and no improvement, and a possible decline, in the 1-Model treatment.....63

Figure 3.3 Experiment 1 final images from participants in the 1-Model and 5-Model treatments. The target image is included at the top for comparison. The columns are chains of participants in the 1-Model treatment. Rows are generations going from top (Generation 1) to bottom (Generation 10). An obvious difference between the two treatments can be seen in the last row.64

Figure 3.4 Mean knot-tying skills over 10 generations for the 1-Model and 5-Model treatments in Experiment 2. Scores rescaled to between 0 and 100, where 100 is a perfect score. The loss of skills is fastest in the first 3 generations and much faster in the 1-Model treatment than in the 5-Model treatment. Generations 4 – 10 suggest different equilibria where the 5-Model treatment has an equilibrium at twice the skill level of the 1-Model equilibrium.....68

Figure 4.1 Flowchart of Experiment Design. The order of the experiments was randomized. We always asked demographic questions at the end, but we asked background measures (not shown) before or after all experiments (also randomized).82

Figure 4.2 Logistic function sigmoid for different values of α (with $c = 0.5$ on left) and different values of c (right). The α parameter determines the curvature of the sigmoid and therefore the strength of the conformist transmission bias. The c parameter determines the inflection point. 88

Figure 4.3 Percentage of decisions that were changed after seeing social information for different number of options. Although there are too few points to be certain about the function that best fits these data, we used a non-linear least squares method to fit to the reciprocal of the number of traits $y = -0.601n + 0.40$, plotted with a grey dashed line. Our choice of fitting the reciprocal of the number of traits is based on the logic underlying the Nakahashi, et al. (2012) model i.e. the probability of selecting the trait at chance is $1/n$89

Figure 4.4 (a) Strength of conformist transmission parameter (α) as a function of number of options (n). The strength of the conformist transmission bias increases with more options. (b) Inflection point of logistic function as a function of number of options. The predicted value based on Nakahashi, et al (2012) is shown as a solid line. The inflection point decreases, but remains higher than the predicted value, indicating an asocial prior.....92

Figure 4.5. Percentage of decisions that were changed after seeing social information for (a) different levels of transmission fidelity, and (b) different question payoff values. Although there are too few points to be certain about the function that best fits these data, we used a non-linear least squares method to fit (a) to a linear model $y = 0.13x + 0.04$, and (b) to a step-function $y = 0.14$ if $x > 0$; $y = 0.11$ if $x = 0$. Fit functions are plotted with a grey dashed line.94

Figure 4.6 (a) Strength of conformist transmission parameter (α) as a function of transmission fidelity. Conformist transmission is strong when fidelity is higher than 60%, but at 60% it's only slightly above unbiased transmission. Strength of conformist transmission parameter (α) as a function of question payoff with (b) all payoff values and (c) \$1 and \$2 averaged to increase sample size for the highest value. The strength of the conformist transmission bias increases with diminishing returns as the payoffs increase.97

Figure 4.7 Density distribution of α conformist transmission values in (a) Experiment 1 and (b) Experiment 2, with α calculated after scaling frequency of options by transmission fidelity. The red line indicates the cut off for conformist transmission with values to the left of this line indicating unbiased social learning. The x-axis is log-scaled. For visualization purposes, we removed some outliers – see Appendix B.5 for figure including these.100

Figure 5.1 Three beta distributions from which values were randomly drawn to simulate individual-level differences and population-level differences in dispositional tendencies toward extraversion and conformity. The symmetrical distribution (long-dashed line) represents individual differences within populations with a moderate mean level of the disposition (equal to the global mean), such as Peru. The right skewed distribution (short-dashed line) represents individual differences within populations with a relatively low mean level of the disposition (approximately 0.5 standard deviations lower than the global mean), such as Morocco. The left skewed distribution (solid line) represents individual differences within populations with a relatively high mean level of the disposition (approximately 0.5 standard deviations higher than the global mean), such as Northern Ireland.122

Figure 5.2 Mean number of opportunities for influence elapsed before majority opinion within a population reached a 2/3 super-majority threshold. (Means computed from 100 simulations for each of the 9 cultural populations. Error bars represent 95% confidence intervals).134

Figure 5.3 Percent of simulations in which a new belief—held initially by just one highly extraverted “lone ideologue”—successfully spread to 50% of the entire population. (Percentage values based on 1000 simulations for each the 9 cultural populations.)140

Figure 5.4 Percent of simulations in which a new belief—held initially by a highly extraverted ideologue along with either 6 disciples or 12 disciples—successfully spread to 50% of the entire population. (Percentage values based on 1000 simulations for each the 9 cultural populations.)144

Figure 6.1. Excerpt from 1869 People’s Magazine article on Muscular Motion.....185

Acknowledgements

As someone who studies cultural transmission, it's difficult not to see the many cultural models I've had in my life. If innovation is mostly cultural recombination, then I have had many sources to draw on and rarely an opportunity to acknowledge them. Early models included Charles Congerton, Wayne Houghton, Chris Everding, Michael Garske, Keith Druery, Trish McGrath, Paul Mead, Wayne Haines, Rosemary O'Neill, and Graeme George. In university, I spent many years in Human Factors and had two amazing mentors—Penny Sanderson and Dave Liu. They introduced me to the *human factor* in the academic enterprise. They gave me a solid foundation in good experimental practices and taught me the importance of remembering the real world, even in basic science. In industry, I had two great supervisors, Tony Williams and Clinton Freeman, whose mentorship made me a better engineer.

The role of more advanced peers in providing a learning gradient should not be underestimated. I am particularly grateful to Daniel Randles, Benjamin Cheung, Aiyana Willard, Wanying Zhao, Maciek Chudek, Sarah Klain, and Will Gervais. I also had exceptional teachers and collaborators in Sue Birch, Michael Doebeli, Patrick Francois, Ara Norenzayan, Vika Savalei, and Ted Slingerland and collaborators in Carl Falk, Kieran Fox, Jordan Levine, Tom Morgan, and Susanne Shultz. I would also like to thank Christoph Hauert and Andrew Whiten, who provided invaluable comments during my dissertation defense.

Some people are lucky enough to have a graduate advisor with whom they are personally and professionally compatible. I had three. Each with different personalities, skills and strengths, intellectual traditions and positions. All highly successful academics who are also just genuinely wonderful people. I couldn't have asked for better cultural models. Steve Heine, Mark Schaller, Joe Henrich—thank you for your guidance and mentorship; it remains an honor working with you.

Like the model presented in Chapter 2, learning goes through two phases. Before my (exceedingly) extended juvenile period, I had 4 important figures in my early childhood. My maternal grandparents, Rama and Saro, and my parents, Clement and Shanthi. Rama and Saro loved and supported me and gave me stability through rocky times. Rama in particular convinced me that I could make the world a better place, and at the very least, I ought to try. I couldn't have asked for a more amazing mother than Shanthi. These four people introduced me to the true diversity in our species and taught me to love and respect that variance. They deliberately instilled in me a love of science, a healthy skepticism of authority, including my own, and a respect for logic and truth over all else.

I would like to thank my amazing team of research assistants and research engineers. I hope your experiences in the lab serve you well in your future careers. I would also like to acknowledge the support of the Vanier Canada Graduate Fellowship, which allowed me to comfortably complete my dissertation. Finally, I would like to thank the remaining countless cultural influences throughout my life. If I were to name you all, these acknowledgements would exceed the remaining chapters.

To my wife, companion, editor, manager, counsel, and now mother to our son, Robert Mathes, to whom this dissertation is also dedicated. Stephanie Salgado, I owe you more than I can express. Without you, none of this would have been possible.

Chapter 1: Introduction

Why are humans so different to all the other animals? When you ask people this question, they usually offer a sensible list: our large brains, use of complex language, our technology and civilization. But underlying this list is a seemingly obvious and unstated assumption: *Humans are just smarter*. But is that really true? Are humans different to other animals simply because we're more intelligent? As it turns out, chimpanzees may have better working memory than we do (Inoue & Matsuzawa, 2007) and even play some economic games closer to the Nash equilibrium – the optimal strategy – when we don't (Martin, Bhui, Bossaerts, Matsuzawa, & Camerer, 2014). When I make these claims in talks, I like to show the audience a video clip of Ayumu, a chimpanzee living in Japan. I ask the audience to play a working memory game against the young chimp: A series of numbers briefly appear on the screen and the audience compete with Ayumu in remembering their locations and then pressing them in order. Ayumu trounces the human audience, pressing all 9 numbers before the humans have even recognized a few¹. The competition is a little unfair. Ayumu has more training in the task and happens to be the genius of the chimp world. But what the video does help convey is that even if humans are smarter, the difference isn't as large as people often assume. At least not at the raw processing level.²

Yet humans are an undeniably remarkable species. Until about 12,000 years ago, our ancestors all lived in small hunter-gatherer groups. About 10,000 years ago, those groups expanded significantly with the advent of agriculture, and by around 6000 years ago, we were living in cities. By 2 and half thousand years ago, Athens has a democracy, 250 years ago the Industrial Revolution begins, 25 years ago we get the World Wide Web, and about 2 and a half years ago, self-driving cars

¹ You can watch the video at muthi.io/ayumu.

² For more discussion on this, see Henrich (forthcoming).

and the Internet of Things³ become a plausible reality. Today, we live in large, interconnected societies of anonymous strangers. So if it wasn't our raw smarts – our working memory, processing, and analytic abilities – then what was it that led us so far from the rest of the animal kingdom? What was it that led us to dominate almost every ecosystem on earth?

In a paper cataloging the differences between humans and chimps, it was in one category that humans truly stood out – our social skills (Herrmann, Call, Hernández-Lloreda, Hare, & Tomasello, 2007). In a comparative study between capuchin monkeys, chimpanzees, and children, Dean, Kendal, Schapiro, Thierry, and Laland (2012) show that it is these social abilities and tendencies that give our species an edge over other smart primates.

My dissertation draws on and contributes to a theory of human behavior called Dual Inheritance Theory or Gene-Culture Coevolution (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981). The theory argues that humans shine in social skills, because humans, unlike the rest of the animal kingdom, have two lines of inheritance. A genetic line, which all species have, and a cultural line, unique to humans. And it is this cultural line, and the interplay between culture and genes, that have allowed our species to adapt to almost every ecology on earth, set up governments, fly to the moon, and build the Internet⁴. Our social skills are a package of abilities and tendencies that have evolved to allow us to acquire that second line of inheritance, the secret to our astonishing achievements. In other words, human smarts are not found in our hardware, they're in our software. Not our brains per se, but what's in them. We are a species hopelessly addicted to information. Information we acquire from peers and those who came before us. We see further, not because we stand on the shoulders of giants, but because we stand on the shoulders of generations upon

³ Internet connected everyday objects doing their job better connected to each other.

⁴ The invention of the Internet has amplified the reach of those social abilities and culture.

generations of accumulated culture - information. Information impossible for even the smartest among us to recreate in a lifetime. And this answer to why humans are so different to all the other animals has huge implications at all levels, from explaining our individual psychology to understanding our emergent social structures and our behavior within them.

1.1 Culture Makes Us Human

The central claim of Dual Inheritance Theory and Cultural Evolution is that the secret of the astonishing success of our species is not hardwired, but acquired. We've adapted to almost every ecology on earth not through raw smarts and individual innovation, but through locally adaptive information accumulated over generations. Boyd and Richerson (1985), two of the original theorists in this area, refer to this information as *culture*. This sometimes causes confusion; in the vernacular, "cultures" are groups of people with distinguishable differences at the population-level i.e. in the Boyd and Richerson (1985) terminology, population-level differences in the distribution of culture. Throughout this dissertation, I will use *culture* to refer to socially transmitted information and *population* to refer to groups of people.

Dual Inheritance Theory or Gene Culture Coevolution is a formal theoretical framework for how our species developed the capacity for culture and particularly cumulative culture, whereby culture accumulates generation by generation to the point where no one could recreate the entire body of culture in a single lifetime. It also describes how culture shapes the environment to which genes adapt in a process of coevolution. In doing so, it makes testable predictions about our physiology and psychology. Cultural Evolution is a subset of Dual Inheritance Theory, which describes how individual psychology playing out at a population-level allows culture to adapt to the local environment, where the local environment including the culture possessed by other individuals in that environment.

1.1.1 Dual Inheritance Theory / Gene-Culture Coevolution & Cultural Evolution

Evolution at its core has 3 prerequisites: (1) variation, (2) transmission or persistence over time, and (3) reduction of that variation. Indeed, the breeder's equation⁵, one of the simplest mathematical representations of evolutionary change, is composed of just these three ingredients. With these three ingredients, any system can evolve (and if that reduction in variation is selective, adapt), whether it's a system of genes, culture, or highly efficient spacecraft antennae created via a genetic algorithm (Hornby, Globus, Linden, & Lohn, 2006). Charles Darwin's great insight was to combine two pieces of culture – artificial selection, which was well understood at the time, and the work of Thomas Malthus. Instead of *artificial* selection by humans, competition for resources could allow populations and individuals to exert selection pressures on each other in a process of *natural* selection. This could explain speciation in a process of “descent with modification”. Evolutionary Theory was mathematically formalized in the early 20th century. By the mid-20th century this formal theory was brought together with other subfields in biology (including genetics) in the Modern Synthesis. In the late 20th century, several theorists (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981; Lumsden & Wilson, 1981) built on formal Evolutionary Theory to explain the specifics of human evolution. They showed how a species might acquire a second cultural line of inheritance to supplement their genetic line, how culture could evolve and enable that species to more quickly adapt to an environment, and how once in place, culture could begin to shape the genetic line. Over time, others added to this body of theory, making further predictions about the psychology, physiology, and population characteristics of such a species. As more cross-cultural and

⁵ $\Delta Z = h^2 S$ where Z is the trait mean in the population, h^2 is heritability (i.e. transmission) and S is the selection differential.

psychological data emerged, these predictions were tested and the theories modified. My dissertation is part of this effort.

1.2 Contributions of Present Research

The present dissertation makes several contributions and addresses several related puzzles. Humans have unusually large brains, 3 times larger than they were a few million years ago and 3 times larger than present-day chimpanzees. And this growth in brain size appears to have taken place to different degrees in several taxa. Many theories have been proposed to explain both this overall trend towards larger brains and the extraordinarily large human brain. These theories have focused on climatic, ecological, and social factors, the most well-known being the Social Brain Hypothesis.

In Chapter 2, I introduce a model of the evolution of large brains that proposes two hypotheses: the (1) Cultural Brain Hypothesis and (2) Cumulative Cultural Brain Hypothesis. The Cultural Brain Hypothesis argues that brains are primarily for the storage and management of adaptive knowledge and that the acquisition of this knowledge has been the primary driver and constraint on the evolution of larger brains. In doing so, it makes several predictions about empirical relationships we should see between brain size, group size, adaptive knowledge, social learning, and the length of the juvenile period. These predictions are consistent with all existing evidence. The Cumulative Cultural Brain Hypothesis argues that these same processes can lead to an autocatalytic take-off between adaptive knowledge and brain size, ratchetting both upwards – the human pathway. The predicted circumstances that trigger this take-off is consistent with other Dual-Inheritance Theory models and empirical evidence, including the evidence presented in Chapters 3 and 4.

The Cumulative Cultural Brain Hypothesis and other treadmill-style models of cultural evolution predict a relationship between sociality and cultural complexity. In contrast, cultural drift style models predict a relationship only between population and cultural complexity. In Chapter 3, I present the results of two experiments that test these two competing theories. In doing so, the experiments also test for the presence of a bias toward learning from people with more adaptive knowledge – a prediction of the Cumulative Cultural Brain Hypothesis⁶. The results of these two experiments support the Cumulative Cultural Brain Hypothesis and other treadmill-style models and also demonstrates the presence of a bias toward social learning from those with more expertise.

In Chapter 4, I present two experiments that test several theories about biased social learning, including the role of transmission fidelity, a key prediction of the Cumulative Cultural Brain Hypothesis. The results of these experiments are consistent with some theories, but not others. The results also reveal that while the use of a conformist transmission bias is ubiquitous, individual differences in cognitive ability still affect the strength and reliance on the bias. Such biases and individual differences have implications at a population level for social networks and the way in which information is transmitted within these networks.

In Chapter 5, I present a theory to explain certain well-studied characteristics of human social networks. The theory is also a useful tool to model human social networks, a requirement for incorporating more realistic social structures into Dual Inheritance Theory models. Chapter 5 also presents a model of the implications of cross-cultural differences in personality and conformist tendencies on the flow of information in society.

⁶ The Cumulative Cultural Brain Hypothesis is not unique in this prediction, but incorporates this bias into a broader evolutionary process.

Chapter 2: Cultural Brain Hypothesis & Cumulative Cultural Brain Hypothesis

In the last few million years, the cranial capacity of hominins dramatically increased, more than tripling in size (Bailey & Geary, 2009; Schoenemann, 2006; Striedter, 2005). This rapid expansion may be part of a gradual and longer-term trend toward larger, more complex brains in many taxa (Dunbar & Shultz, 2007a; Roth & Dicke, 2005; Shultz & Dunbar, 2010a; Striedter, 2005). These patterns are puzzling since brain tissue is energetically expensive (Aiello & Wheeler, 1995; Foley, Lee, Widdowson, Knight, & Jonxis, 1991; Isler & Van Schaik, 2006; Kotrschal et al., 2013). Efforts to understand the evolutionary forces driving brain expansion, both in the human lineages, more broadly with primates, and across taxa, have focused on climatic, ecological, and social factors (see Bailey & Geary, 2009; Dunbar, 2003; Schoenemann, 2006; Striedter, 2005; van Schaik & Burkart, 2011). One particularly popular explanation is the Social Brain Hypothesis (Dunbar, 1998), which argues that brains have primarily evolved for dealing with the complexities of life in larger groups (keeping track of individuals, Machiavellian reasoning, and so on). The primary evidence in favor of the Social Brain Hypothesis is the empirical relationship that's been shown between social group size in primates (or other measures of sociality in other species) and some measure of brain size (these measures of brain size are typically highly correlated; see Dunbar, 2009). There have been some attempts to model the mechanisms underlying this relationship (e.g. Dávid-Barrett & Dunbar, 2013; Gavrilets & Vose, 2006). Here, we develop a model and evolutionary simulation of two hypotheses, the Cultural Brain Hypothesis (Pradhan, Tennie, & van Schaik, 2012; van Schaik & Burkart, 2011; van Schaik, Isler, & Burkart, 2012; Whiten & Van Schaik, 2007) and Cumulative Cultural Brain Hypothesis (Boyd & Richerson, 1996; Henrich & McElreath, 2003; Herrmann et al., 2007; Heyes, 2012). The broad idea is that rather than for specifically dealing with the complexities of life in larger groups, brains have been selected for their ability to store and manage information,

which may be acquired via various forms of individual (asocial) or social learning. This information is locally adaptive on average and is plausibly related to information about other group members and competitive strategies, but may also be related to solving problems such as finding and processing food, avoiding predators, making tools, and locating water.

The presence of social learning has now been confirmed in many species, including many primate species (Hoppitt & Laland, 2013). Given this, the Cultural Brain Hypothesis proposes that the same underlying selective process that led to widespread social learning may also explain the correlations observed across species in variables related to brain size, group size, social learning, innovation, and life history. It may also explain why brains have expanded more in some lineages than others (Dunbar & Shultz, 2007a; van Schaik et al., 2012). Building on this, the Cumulative Cultural Brain Hypothesis proposes that under some conditions social learning may cause a body of adaptive information to accumulate over generations. This accumulating body of information may lead to selection for brains better at social learning, and storing and managing this adaptive knowledge. Larger brains, better at social learning, could then permit the body of adaptive knowledge to expand even more, creating an autocatalytic feedback loop that drives up both brain size and adaptive knowledge in a culture-gene co-evolutionary duet – the uniquely human pathway.

This approach seats humans within the broader evolutionary story. The same mechanisms that lead to widespread social learning can also lead to the human pathway in a very specific set of circumstances. However, while the Cultural Brain Hypothesis and Cumulative Cultural Brain Hypothesis are related, we keep them distinct for two reasons. First, the cumulative culture-gene co-evolutionary process produces cultural products, like sophisticated multi-part tools, that no single individual could figure out in their lifetime (despite having a big brain capable of potent individual learning). The evolution of a second line of inheritance – culture – is a qualitative shift in the

evolutionary process. Second, it's possible that either of these hypotheses could hold without the other fitting the evidence.

To develop these hypotheses, our simulation explores the interaction and coevolution of (1) learned adaptive knowledge and (2) genetic influences on brain size (storage capacity), asocial learning, social learning, and an extended juvenile period with the potential for payoff-biased oblique social learning. We explicitly model population growth and carrying capacity alongside genes and learning in order to theorize potential relationships between group size and other parameters, like brain size and adaptive knowledge, and also to examine the effects of sociality on the co-evolutionary process through two different parameters. We assume carrying capacity is increased by the possession of adaptive knowledge (e.g. can access more calories, better avoid predators). Our model incorporates ecological factors and phylogenetic constraints by considering different relationships between birth/death rates and both brain size and adaptive knowledge.

In constructing this model, we make the following three basic assumptions:

1. Larger and more complex brains are more costly than less complex brains because they require more calories, are harder to birth, take longer to develop, and have organizational challenges. Therefore, *ceteris paribus*, increasing brain size decreases an organism's fitness. For simplicity, we assume that brain size, complexity, and organization (e.g. neuronal density) is captured by a single state variable, which we will refer to as "size".
2. A larger brain correlates with an increased capacity and/or complexity that allows for the storage and management of more adaptive knowledge. Adaptive knowledge relates to locating food, food preparation (detoxification, increased calorie release),

hunting techniques, identifying medicinal plants, tool-making skills, weather-resistant clothing, and so on.

3. More adaptive knowledge increases an organism's fitness either by increasing its number of offspring compared to conspecifics or by reducing its probability of dying before reproduction, or both. Adaptive knowledge can be acquired asocially, through experience and causal reasoning, or socially, by learning from others.

Our model makes predictions about the relationships between brain size, group size, adaptive knowledge, social learning, mating structure, and the juvenile period. Some of these relationships have been tested in the empirical literature and our model's predictions are supported by these data (see Section 2.3.2). Specifically, several authors have shown positive relationships (notably in primates) between (1) brain size and social group size (Social Brain Hypothesis; Barton, 1996; Dunbar & Shultz, 2007a; Dunbar, 1998), (2) brain size and social learning (Lefebvre, 2013; Reader & Laland, 2002), (3) brain size and length of juvenile period (Charvet & Finlay, 2012; Isler & van Schaik, 2009; Joffe, 1997; Walker, Burger, Wagner, & Von Rueden, 2006), (4) group size and the length of the juvenile period (Joffe, 1997), and (5) group size or other measures of sociality and toolkit size (Lind & Lindenfors, 2010; van Schaik et al., 2003; Whiten & Van Schaik, 2007). Our model predicts all these empirical relationships, uniting these disparate findings under a single, more general theory: The Cultural Brain Hypothesis. In addition, we find that differences among taxa (e.g. Shultz & Dunbar, 2010a) in these relationships can be accounted for by λ , which moderates the effect of adaptive knowledge on the death rate in our model. The λ parameter can be interpreted as being part of the resource richness of the ecology. Richer ecologies offer more "bang for buck", for example, more calories unlocked for less knowledge, allowing individuals to better offset the size of their brains. Higher λ suggest a richer ecology.

The dynamics of our model also reveal the ecological conditions and evolved psychology most likely to lead to the realm of cumulative cultural evolution, the pathway to modern humans. We call these predictions the Cumulative Cultural Brain Hypothesis. Our model indicates the following pathway. Under some conditions, brains expand to improve asocial learning and thereby create more adaptive knowledge. This pool of adaptive knowledge leads to selection favoring an immense reliance on social learning, with selective oblique transmission, allowing individuals to exploit this pool of knowledge. Rogers' (1988) paradox (whereby social learners benefit from taking advantage of asocial learner's knowledge, but do not themselves generate new knowledge) is solved by selective oblique social learning transmitting accidental innovations to the next generation. Under some conditions, an interaction between brain size, adaptive knowledge, and sociality (deme size and interconnectedness) emerge, creating a feedback loop that drive all three - the beginning of cumulative cultural evolution.

2.1 Model Description

To explore the culture-gene co-evolutionary dynamics, we constructed an evolutionary simulation in which individuals are born, learn asocially or socially from their parent, potentially update by asocial learning or by socially learning from more successful members of their group during an extended juvenile period, migrate between demes, and die or survive based on their brain size and adaptive knowledge. Individuals who survive this process give birth to the next generation. Because we use a haploid population and are mainly interested in the effects of natural selection and learning, we will ignore non-selective forces such as sex, gene recombination, epistasis, and dominance.

The simulation⁷ begins with 50 demes, each with a population of 10 individuals. Throughout the simulation, the number of demes was fixed at 50. In early iterations of the model, we explored increasing the number of demes to 100 for some of the parameter space and found no significant impact on the results. Our starting population of 10 individuals is roughly equivalent to a real population of 40 individuals, assuming two sexes and one child per parent (4 x 10). As a reference, mean group size in modern primates ranges from 1 to 70 (Dunbar, 1992).

Each individual i in deme j has a brain of size b_{ij} with a fitness cost that increases with increasing brain size. Adaptive knowledge is represented by a_{ij} , where $0 \leq a_{ij} \leq b_{ij}$. Increasing adaptive knowledge can mitigate the selection cost of a larger brain.

Our simulations begin with individuals who have no adaptive knowledge, but the ability to fill their $b_{ij} = 1.0$ sized brains with adaptive knowledge through asocial and/or social learning. To explore the idea that juvenile periods extend the time permitted for learning, we have included two stages of learning. In both learning stages, the probability of using social learning rather than asocial learning is determined by an evolving *social learning probability* variable (s_{ij}). In the results reported here, the social learning probability variable is initially set at one (i.e. at the beginning of the simulation all individuals are social learners). We ran this version to show that under all conditions social learners are invaded by asocial learners (see Figure 2.11b); social learning is only valuable when sufficient adaptive knowledge exists in the group. The starting condition affect the initial evolution of s , but not equilibrium values.

Asocial learning allows for the acquisition of adaptive knowledge independent of the adaptive knowledge possessed by other individuals. In contrast, social learning allows for vertical acquisition

⁷ The simulation was programmed in C++ and is available on request. We tested for bugs using a suite of tests written using the Google C++ Testing Framework.

of adaptive knowledge possessed by the genetic parent in the first learning stage or oblique acquisition from more knowledgeable members of the deme (from the parent generation) in the second learning stage. The tendency to learn from models other than the genetic parent is determined by a genetically evolving *oblique learning probability* variable (v_{ij}). Thus the probability of engaging in oblique learning is an indication of the length of the juvenile period. In the second stage of learning, if an individual tries to use social learning, but does not use oblique learning, no learning takes place beyond the first stage, creating an initial advantage for asocial learning. The ability to select a model with more adaptive knowledge when oblique learning is determined by an evolving *oblique learning bias* variable (l_{ij}). In the next section, we discuss each stage of this lifecycle in more detail.

2.1.1 The Lifecycle

Individuals go through four distinct life stages (Figure 2.1): Individuals (1) are born with genetic traits similar to their parents, with some mutation, (2A) learn adaptive knowledge socially from their parents or through asocial learning independent of their parents, (2B) go through a second stage of learning adaptive knowledge through asocial or oblique social learning, (3) migrate between demes, and (4) die or survive to reproduce the next generation. As detailed below, we model both fecundity and viability selection as separate processes.

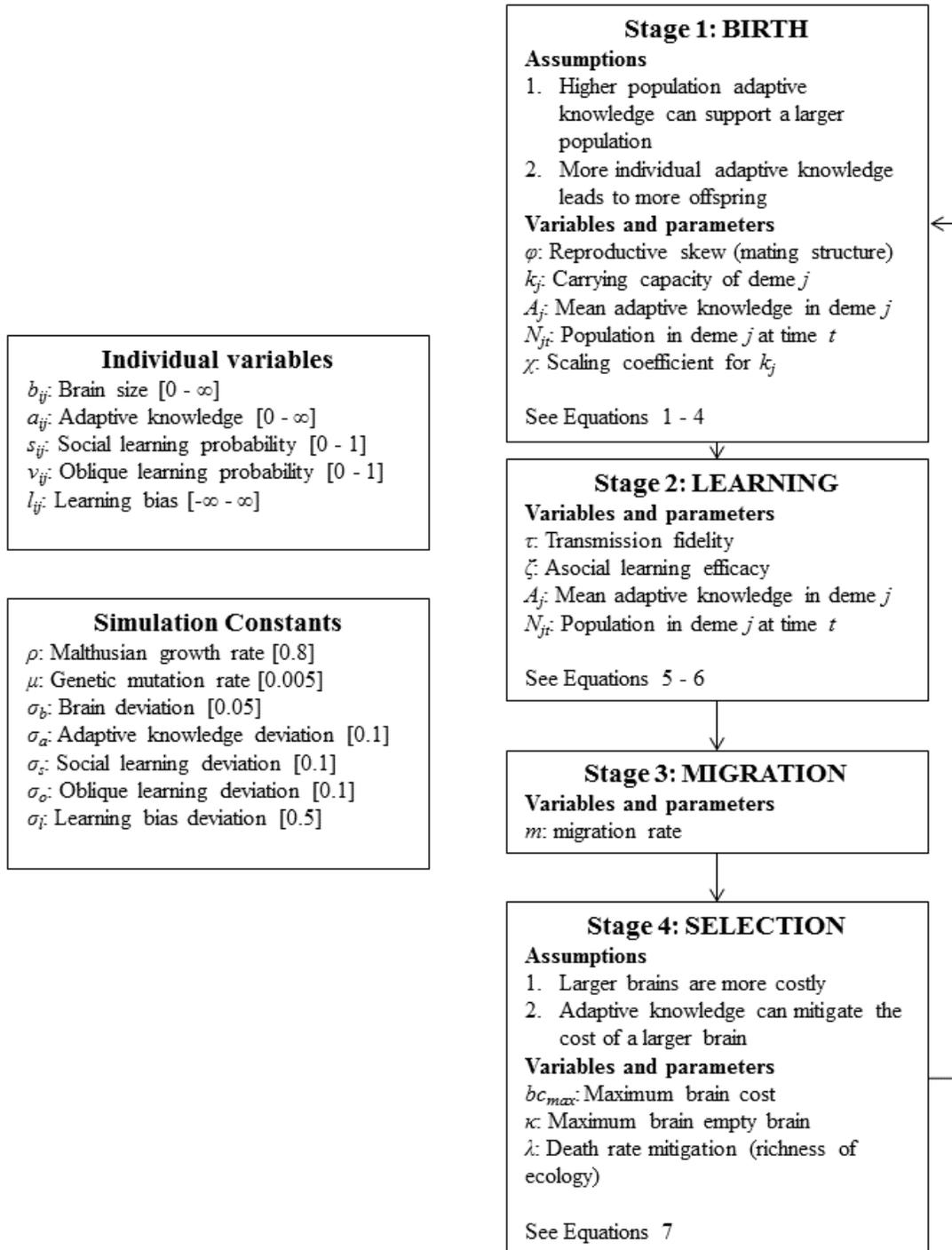


Figure 2.1. Simulation variables, parameters, and flowchart. Four life stages for all individuals.

2.1.2 Stage 1: The Birth Stage

In the birth stage, the individuals who survive the selection stage (Stage 4) give birth to the next generation.

2.1.2.1 Adaptive Knowledge and the Number of Children

We assume that demes with greater mean adaptive knowledge can sustain a larger population. We formalized this assumption in (1) by linking the carrying capacity of a deme (k_j) to the mean adaptive knowledge of the individuals in the deme (A_j). The carrying capacity of the current generation (k_j) is then used to calculate the total expected number of children in the next generation ($N_{j_{t+1}}$) using the discrete logistic growth function in (2) where ρ is the generational growth rate. Pianka (2011) estimates a maximum intrinsic rate of growth of 2.1 per generation for *homo sapiens* assuming a 20 year generation length. Since this is a maximum, we use a more conservative value ($\rho = 0.8$).

$$k_j = \chi A_j + N_{j_0} \quad (1)$$

$$N_{j_{t+1}} = \frac{e^{\rho N_{jt}}}{1 + \left((e^{\rho} - 1) \frac{N_{jt}}{k_j} \right)} \quad (2)$$

A potential parent's (i_j) probability of giving birth (p_{ij}) is given by their sigmoid transformed adaptive knowledge value (3) as a fraction of the sum of all transformed adaptive knowledge values of individuals in the deme (4). We assume that more individual adaptive knowledge (a_{ij}) is associated with increased relative fertility. The actual number of children n_{ij} for each parent is drawn from a binomial distribution $B(N_{j_{t+1}}, p_{ij})$. By drawing these values from a binomial distribution, the sum of expected values for all parents is $N_{j_{t+1}}$ (i.e. $N_{j_{t+1}} = \sum_{i=1}^{N_j} E[B(N_{j_{t+1}}, p_{ij})] = E[\sum_{i=1}^{N_j} n_{ij}]$). Children are endowed with genetic characteristics similar to their parents.

$$a_{ij}^T = \frac{1}{1 + e^{-\varphi(a_{ij} - A_j)}} \quad (3)$$

$$p_{ij} = \frac{a_{ij}^T}{\sum_{i=1}^{N_j} a_{ij}^T} \quad (4)$$

This parameterization allows us to study the importance of fecundity selection. For example, we can turn-off fecundity selection entirely by setting $\varphi = 0$: A world with no reproductive skew; all potential parents have the same probability of giving birth. The more we turn up φ , the more we have a winner take all world, where to win one has to acquire adaptive knowledge. This is crucial in thinking about how, for example, our culture-gene co-evolutionary process is influenced by social organization and mating structures that create high reproductive skew. One mechanism underlying reproductive skew is mating structure. A perfectly monogamous pair-bonded society with no differential selection at the birthing stage would have $\varphi = 0$. Increasing φ allows for an increase in polygyny from “monogamish” (mostly pair-bonded) cooperative breeding societies at low values of φ to a highly polygynous winner-takes all societies where males with the most adaptive knowledge have significantly more children. Our model suggests that in more polygynous societies, where selection is high, variation is reduced. This allows for the initial rapid evolution of larger brains, but with little or no variation, populations are unable to use social learning to increase their adaptive knowledge and are more likely to go extinct. At the other extreme, evolutionary forces are quashed when $\varphi = 0$. Social learning and the advent of culture-gene coevolution are more likely to occur in values of φ that correspond to pair-bonded and cooperative breeding societies. Of course, some argue that culture supports or is responsible for such mating structures in humans. We have not endogenized φ .

2.1.2.2 Genetic Transmission and Mutation

Individuals acquire four genetic traits from their parents - their brain size (b'_{ij}), social learning probability (s'_{ij}), oblique learning probability (v'_{ij}), and oblique learning bias (l'_{ij}). For each trait, newborn individuals have a $1 - \mu$ probability of having the same value as their parents ($b_{ij}, s_{ij}, v_{ij}, l_{ij}$). If a mutation takes place, new values are drawn from a normal distribution with a mean of their parent value and a standard deviation σ_s for s'_{ij} , σ_v for v'_{ij} , σ_l for l , and σ_b for b'_{ij} and $\sigma_b b_{ij}$ for b'_{ij} . The standard deviation of s'_{ij} and v'_{ij} are not scaled by the mean since they are probabilities and therefore bounded $[0,1]$. Although l'_{ij} is not bounded, we do not scale the standard deviation by the mean, because small changes in l'_{ij} have a large effect on learning bias due to the sigmoid function. Once children have been endowed with genetic characteristics, they then acquire adaptive knowledge based on these characteristics.

2.1.3 Stage 2: Learning

Asocially learned adaptive knowledge values (a'_{ij}) are drawn from a normal distribution based on an individual's brain size: $N(\zeta b'_{ij}, \sigma_a \zeta b'_{ij})$. The idea here is that bigger brains will be better at solving novel problems, and figuring stuff out (Deaner, Isler, Burkart, & van Schaik, 2007; Sol, Bacher, Reader, & Lefebvre, 2008). But, many other factors will influence individuals' ability to use that brain, such as constraints on time (for trial and error learning) or energy. These constraints are captured by ζ .

Socially learned adaptive knowledge values are drawn from a similar normal distribution, but with a mean of the model's (t) adaptive knowledge value scaled by transmission fidelity (τ):

$N(\tau a_{tj}, \sigma_a \tau a_{tj})$. Here τ could include individuals' cognitive abilities, but we focus on the social elements of transmission fidelity. Greater social tolerance, more interactions or opportunities for

interaction, and some passive or active teaching by models will increase transmission fidelity (Dean et al., 2012; Whiten & Erdal, 2012). With greater complexity, this effect could be broken down into constraints (parameters) and endogenous state variables (e.g. genes for sociality), but here we capture all this with τ .

For both asocial and social learning, an agent's adaptive knowledge may not exceed their brain size. But, compared to social learning, asocial learning enables the immediate acquisition of adaptive knowledge based on one's own brain size. Social learning is dependent on the adaptive knowledge possessed by parents, or those in the parent's generation within the same deme if selection extends the learning phrase through a juvenile period.

2.1.3.1 Stage 2A: Vertical Transmission or Asocial Learning

In Stage 2A, newborn individuals i'_j can socially acquire adaptive knowledge from their parent i with probability s'_{ij} . If newborns do not learn from their parents ($1 - s'_{ij}$), they instead learn asocially.

2.1.3.2 Stage 2B: Asocial learning, Oblique learning, or No update

In Stage 2B, individuals i'_j may update their adaptive knowledge through asocial learning with probability $(1 - s'_{ij})$ in the same manner as Stage 2A or obliquely from non-parents with probability $s'_{ij}v'_{ij}$. Individuals who do not asocially learn nor obliquely learn do no further learning. This allows us to study conditions under which oblique learning emerge during this extended learning period. Crucially, oblique learning has to out compete a second round of asocial learning.

We adjust the strength of the relationship between a potential model's (t) adaptive knowledge and their likelihood of being modeled using the learner's l'_{ij} variable in the sigmoid transformation function (5). A potential model's (t_j) probability of being selected (p_{tj}) is given by (6). Both asocial

and social learning only updates adaptive knowledge values if these values are larger than those acquired during the first stage of learning, Stage 2A.

$$a_{tj}^T = \frac{1}{1 + e^{-l'_{ij}(a_{tj} - A_j)}} \quad (5)$$

$$p_{tj} = \frac{a_{tj}^T}{\sum_{t=1}^{N_j} a_{tj}^T} \quad (6)$$

Note, since we are interested in the evolution of social learning, we specifically stacked the deck somewhat against social learning. Individuals have a $s'_{ij} - s'_{ij}v'_{ij}$ chance of not doing *any* learning during Stage 2B, which creates an initial disadvantage for social learning since any selection for social learning in Stage 2A risks missing out on a second round of asocial learning in Stage 2B.

2.1.4 Stage 3: Migration

Individuals migrate to a randomly chosen deme (not including their own) with probability m . All demes have the same probability of immigration. Individuals retain their adaptive knowledge and genetic traits. There is no selection during migration; all individuals survive the journey.

2.1.5 Stage 4: Selection Based on Brain Size and Adaptive Knowledge

We formalized the assumption that larger, more complex brains are also more costly using both a quadratic and exponential function to link brain size to maximum death rate (c_{max}). These functions capture the idea that the costs of large brains escalate non-linearly with size. Here we focus only on the quadratic function $c_{max} = \frac{b'_{ij}{}^2}{\kappa^2}$, as a comparison of the two functions revealed no important qualitative differences.

To formalize the assumption that individuals with more adaptive knowledge are less likely to die *ceteris paribus*, we use the negative exponential function in (7). The λ parameter in (7) was varied between simulations and was used to determine the extent to which adaptive knowledge can offset the costs of brain size, where $\lambda = 0$ indicates no offset. The λ parameter can be interpreted as how

much adaptive knowledge one requires to unlock fitness-enhancing advantages. For example, in a calorie-rich environment where only a little smarts are required to access calories (e.g. remembering food locations), λ would be high. Conversely, in a calorie-poor environment where a lot of smarts are required to access fewer calories (e.g. food needs significant preparation before safe consumption), λ would be low. This relationship between λ and death rate (d) is graphed in Figure 2.2 for different λ values.

$$d = c_{max} e^{-\lambda \frac{a'_{ij}}{c_{ij}}} \quad (7)$$

This function captures the idea that the increasing costs of big brains can be offset by more adaptive knowledge. The choice of setting the maximum empty brain size $\kappa = 100$ was somewhat arbitrary, but allowed for a reasonable sized brain to see a range of evolutionary behavior.

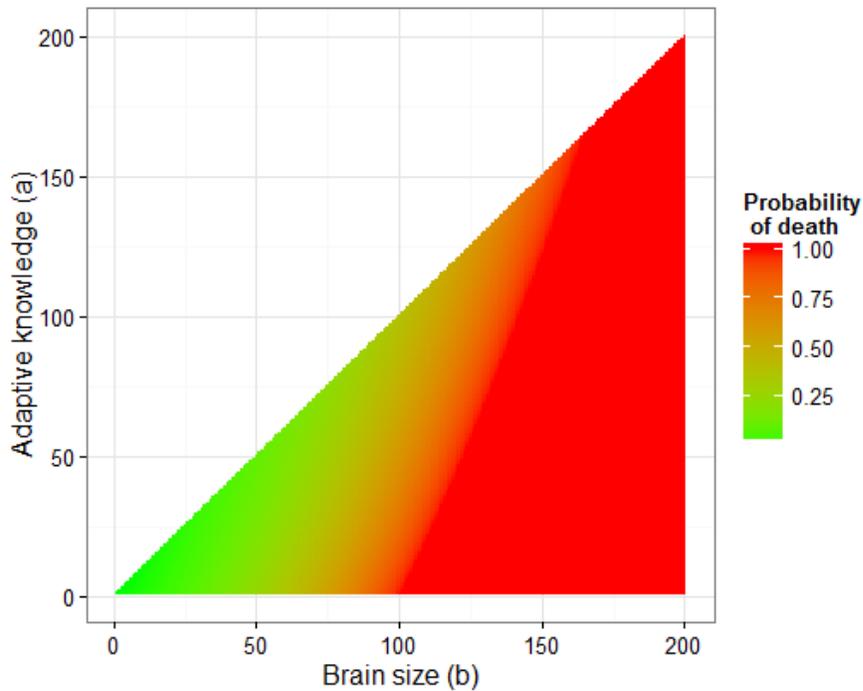


Figure 2.2. Fitness cost as measured by death rate plotted for different brain sizes and adaptive knowledge. Red indicates a higher death rate.

2.1.6 Summary

These basic assumptions generate conflicting selection pressures for (1) more adaptive knowledge and (2) smaller brains. Under some conditions, the cost of having a larger brain is offset by the increased amount of adaptive knowledge a larger brain can handle. If adaptive knowledge were freely available, there would be no constraint on the co-evolution of brains and adaptive knowledge; both would ratchet upward. In general, three related constraints prevent this from happening:

1. Adaptive knowledge does not always exist in the environment to fill your larger brain.

2. Your larger brain without adaptive knowledge is costly without benefit. This is especially true for social learners with brains larger than their parents, since this additional brain space cannot be utilized.
3. Increases in brain size show diminishing returns; brain costs increase at a greater than linear rate.

We simulated a range of space within each parameter set for low, middle, and high values of other parameters for which we found interactions and realistic values of all other parameters. The range for each parameter was as follows: $\varphi[0.0,1.0]$, $\tau[0.75,1.0]$, $\zeta[0.1,0.9]$, $m[0.0,0.2]$, and $\lambda[0.0,2.0]$.

To give our populations enough time to evolve, we ran our simulation for 200 000 generations. Assuming 20-25 years per generation, this represents 4-5 million years of evolution, approximately the time since the hominin split from chimpanzees. With a few exceptions, this guarantees that our genetically evolved variables have hit a quasi-equilibrium. To account for stochastic variation in simulation outcomes, we performed 5 iterations per set of unique parameters and averaged the results across these. Learning bias l did not reach equilibrium, however, we would not expect it to do so since higher l values continue to provide an advantage in selecting models, such that l should slow down but continue to approach ∞ . In our model, l is a one dimensional state variable that captures better and worse ability to select models, but of course in the real world, there are a range of strategies and biases that have evolved to solve the problem of selecting models with more adaptive knowledge (for a list of such biases and strategies, see Rendell et al., 2010).

2.2 Theory

To organize our presentation, we first consider the Cultural Brain Hypothesis. The Cultural Brain Hypothesis makes several predictions about the relationships between brain size, group size,

adaptive knowledge, social learning, and the length of the juvenile period. Many of these relationships have been measured across real species. To test the predictions of the Cultural Brain Hypothesis, we treat the quasi-equilibrium outcomes of each of our simulation runs as species that have evolved, with state variables representing the characteristics of each species. We then look across all the parameter values we explored, which represent different environments or phylogenetic constraints, and examine the relationships between our state variables in the same manner in which they have been tested in the empirical literature. It bears emphasis that all the outcome variables we will be looking at were coevolving individual-level state variables with no direct formulaic relationships. Of course, there are indirect relationships, and these are the basis of the Cultural Brain Hypothesis theory and the focus of our investigation.

Next, we focus on the Cumulative Cultural Brain Hypothesis and examine the conditions that favor substantial amounts of cumulative cultural evolution. The goal here is to understand the conditions under which the interaction between social learning, brain size, group size, sociality and life history may generate the kind of auto-catalytic takeoff required to explain the last two million years of human evolution. Again, where possible, we test these theories against the empirical literature.

2.3 Cultural Brain Hypothesis

The relationship between many variables of interest have been tested in the empirical literature. Following the same analytic approach used for each of these empirical relationships, the Cultural Brain Hypothesis predicts a positive relationship between brain size and group size, brain size and social learning, brain size and the length of the juvenile period, group size and the length of the juvenile period, and group size and toolkit size. Each of these predictions are borne out by the empirical literature. Note that: (1) The exact correlations between each of these empirical variables

does not have to be the same as our theoretical predictions for the prediction to be borne out – these models make qualitative predictions about the relative size of relationships (i.e. some relationships are stronger than others). (2) We did not select parameter values to fit the empirical literature, but rather a range of values that encompass realistic values. We tested a range of values for each parameter and report this range and the effect of each parameter here.

2.3.1 Predictions

We have 5 parameters in our model: Reproductive skew (mating structure; φ), transmission fidelity (τ), asocial learning efficacy (ζ), migration rate (m), and richness of the ecology (λ). Each represents different ecological and phylogenetic constraints. Under different combinations of these conditions, we find that at least two regimes emerge: (1) Species which mostly rely on asocial learning and (2) species which mostly rely on social learning. We can see these regimes in a histogram of mean social learning values (s) across all simulations (Figure 2.3). A k-cluster analysis on the social learning mean (s) confirms that the threshold between these regimes is 50%. Note that the relative count size of the two regimes is a reflection of the range of parameters we chose rather than a reflection of the world. We chose parameters that encompassed the range of realistic values, but allowed us to investigate the threshold between asocial and social learning (e.g. transmission fidelity values greater than 75% rather than from 0% to 100%). Under some conditions, a species which relies on social learning can enter into the realm of cumulative cultural evolution. We discuss the conditions when this transition is most likely to take place in Section 2.3.3 on the Cumulative Cultural Brain Hypothesis. The relationships between equilibrium state variable values differ considerably between these two regimes and so we analyze them separately.

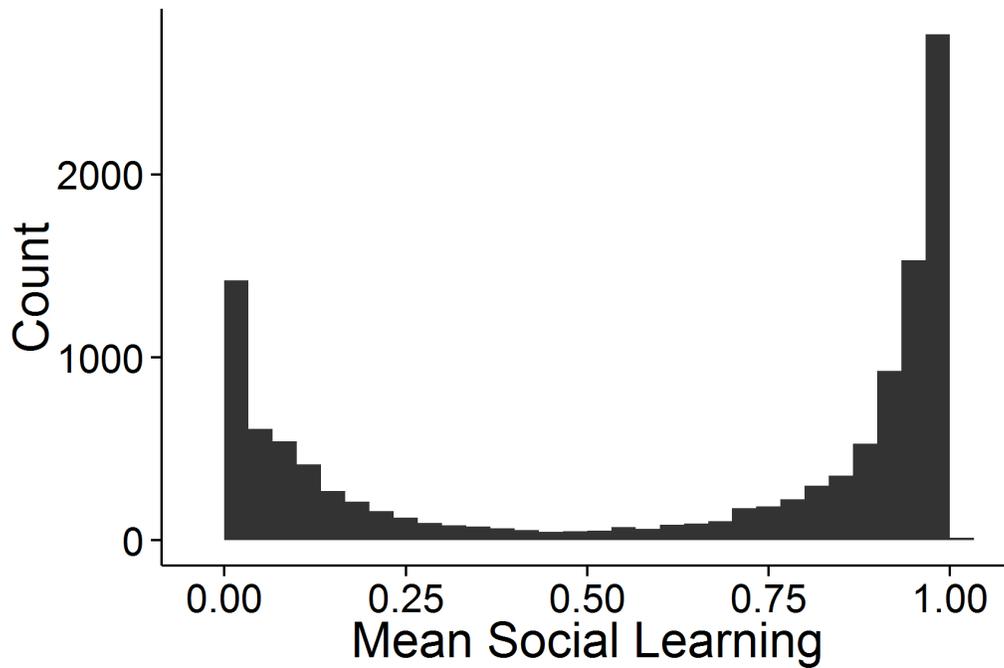


Figure 2.3. Histogram of mean social learning probability (s) across all simulations. Note that the relative count size of the two regimes is a reflection of the range of parameters we chose rather than a reflection of the world. We chose parameters that encompassed a realistic range, but allowed us to investigate the threshold between asocial and social learning (e.g. transmission fidelity values greater than 75% rather than 0% to 100%).

Of the 5 parameters, richness of the ecology (λ) affects the maximum possible brain size. Therefore, in Table 2.1 and

Table 2.2 below, we report the value of these relationships for a particular value of $\lambda = 1$ so as to illustrate the relative relationships between these values. We analyze these relationships in a manner consistent with the empirical literature so as to test our model. We discuss the processes underlying these relationships, including the effects of λ in Section 2.3.3.

Table 2.1. Correlations between log mean brain size, log mean adaptive knowledge, log mean group size, mean social learning, and mean juvenile period across the entire parameter space. The table has been color coded from white ($r = 0$) to blue ($r = 1$) for ease of comprehension.

Mangel & Clark, 1988	Brains $\log(\bar{b})$	Adaptive Knowledge $\log(\bar{a})$	Group Size $\log(\bar{N})$	Social Learning $\log(3 + \bar{N}s)$	Juvenile Period \bar{sv}
$\log(\bar{b})$	1				
$\log(\bar{a})$	0.80	1			
$\log(\bar{N})$	0.50	0.82	1		
$\log(3 + \bar{N}s)$	0.19	0.61	0.87	1	
\bar{sv}	0.02	0.48	0.66	0.91	1

Table 2.2. Correlations between log mean brain size, log mean adaptive knowledge, log mean group size, mean social learning, and mean juvenile period. Primarily asocial learners ($s < .5$) are in the bottom triangle and primarily social learners ($s > .5$) are in the top triangle. The table has been color coded from red ($r = -1$) to blue ($r = 1$) for ease of comprehension.

	Brains $\log(\bar{b})$	Adaptive Knowledge $\log(\bar{a})$	Group Size $\log(\bar{N})$	Social Learning $\log(3 + \bar{N}s)$	Juvenile Period \bar{sv}
$\log(\bar{b})$	1	0.98	0.81	0.82	0.48
$\log(\bar{a})$	0.87	1	0.84	0.86	0.47
$\log(\bar{N})$	0.41	0.71	1	0.99	0.36
$\log(3 + \bar{N}s)$	-0.53	-0.19	0.37	1	0.44
\bar{sv}	-0.59	-0.45	-0.19	0.62	1

2.3.2 Testing Predictions

We have analyzed the output of our simulation using the same approaches taken in the empirical literature. That is, in Table 2.1 and

Table 2.2, we have treated our quasi-equilibrium outcomes as the characteristics of extant species and analyzed them as if they were empirical data. In doing so, we can compare our predictions to

real data, but of course in reality, these are qualitative predictions; our simulations cannot capture all aspects of the evolutionary process. Instead, we hope that they capture the minimum elements needed to explain a wide array of empirical findings. Therefore, at a minimum, we expect these predictions to be consistent with patterns found in the empirical literature. Assuming we can predict these empirical patterns, the real value of the model is in explaining why these patterns exist. The model also makes additional predictions about thus far untested relationships.

2.3.2.1 Brain Size and Group Size

The relationship between brain size and group size is the basis for the Social Brain Hypothesis – the idea that the evolution of large brains is primarily driven by the complexities of life in larger groups. In our model, there is no direct effect of brain size on group size or group size on brain size. Instead, brains are assumed to be more general, allowing for the storage and management of more information. Our model indicates that among species that primarily rely on social learning, the relationship between brain size and group size is consistently positive ranging somewhere between $r = .52$ and $r = .88$ (depending on λ), the lower correlation being when adaptive knowledge provides no cost offset to large brains (i.e. λ is low). In contrast, our model predicts that among taxa with that rely more on asocial learning, the relationship is much weaker, ranging from $r = .25$ to $r = .71$. See Figure 2.4.

The empirical literature has established a strong positive relationship between brain size and group size in primates (Barton, 1996; Dunbar, 1998), but not in other taxa (see Dunbar, 2003; Dunbar & Shultz, 2007b; Pérez-Barbería, Shultz, & Dunbar, 2007). In primates, the correlation between relative neocortex size and group size is somewhere between $r = .48$ to $r = .61$ (Barton, 1996). In support of the Social Brain Hypothesis, researchers have noted that in other taxa, brain size correlates with measure of sociality (e.g. among non-primate mammals) and with mating

structure (e.g. among birds). However, as Dunbar and Shultz (2007) admit, why group size only correlates with brain size in some taxa and not others remains a mystery. The Cultural Brain Hypothesis offers an explanation.

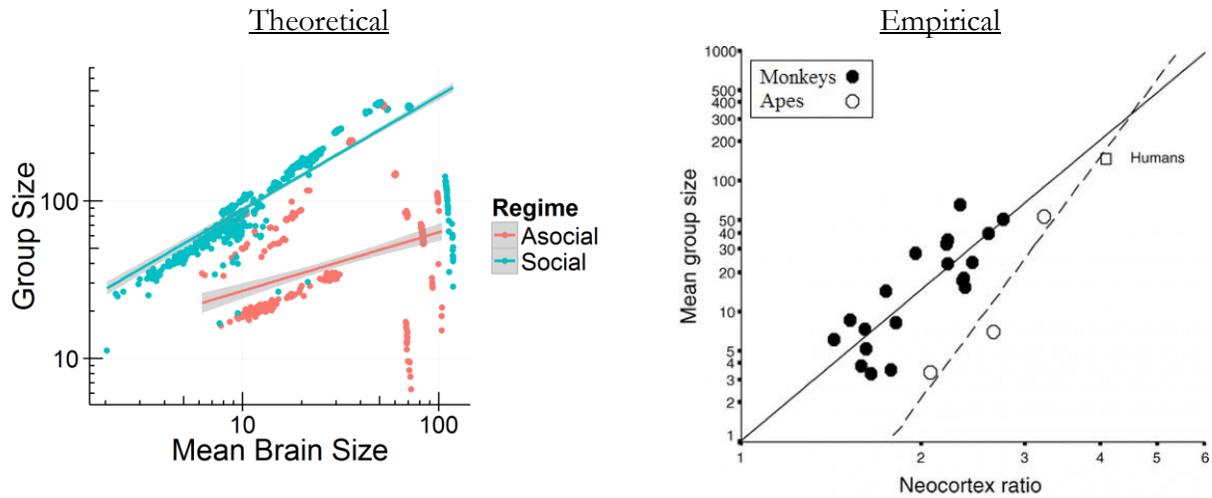


Figure 2.4. Example of correlations between brain size and group size for regimes that emerge in our model (left). Correlation between brain size and group size in the empirical literature (right; Barton, 1996).

The Cultural Brain Hypothesis assumes that larger brains are better at storing and managing adaptive knowledge. There are two pathways to acquire that knowledge – asocial learning and social learning. Groups with higher mean adaptive knowledge have a higher carrying capacity. This relationship exists regardless of whether that knowledge is acquired socially or asocially, thus taxa more reliant on asocial learning generally also have a small to moderate relationship between brain size and group size in our model (varying on the effectiveness of asocial learning). For taxa more reliant on social learning, larger groups also have more adaptive knowledge to exploit, raising the mean adaptive knowledge of the group and therefore the carrying capacity. Thus there is a strong relationship between brain size and group size among more social learning taxa in our model. This

logic of the Cultural Brain Hypothesis also predicts that brain size should correlate with social learning.

2.3.2.2 Brain Size and Social Learning

The Cultural Brain Hypothesis predicts a positive relationship between brain size and social learning among species that engage in social learning. Social learners in a knowledge filled environment are more effectively able to fill their large brains than asocial learners, using that adaptive knowledge to pay for the calorific cost of larger brains and creating a pressure for larger brains to store and manage that knowledge. The empirical measures of social learning represent counts of observations of social learning, showing a correlation of $r = .69, p < .001$ for primates (Reader and Laland, 2002; Lefebvre, 2013). To better match our social learning probability, s , to the empirically available results, we assumed that simulated species with larger populations and higher s values would generate greater numbers of observations (linearly). Thus, we multiplied s by mean group size (N), and then following the empirical work, added 3, and took the natural log (Reader and Laland, 2002). Our model indicates that among species that primarily rely on social learning, the relationship between brain size and group size is consistently positive ranging somewhere between $r = .56$ and $r = .89$, the lower correlation when λ is lower. Our model predicts that among species that primarily rely on asocial learning, the relationship is negative. Most species remain small brained, but those that do acquire a large brain via asocial learning, do so at the expense of social learning abilities. See Figure 2.5. While such a situation may be theoretically possible – highly intelligent species relying on their own wit alone – we know of no such species in the real world and thus the higher asocial learning efficacy values (ζ) that lead to such a situation may be unrealistic. We discuss this in Section 2.3.3.

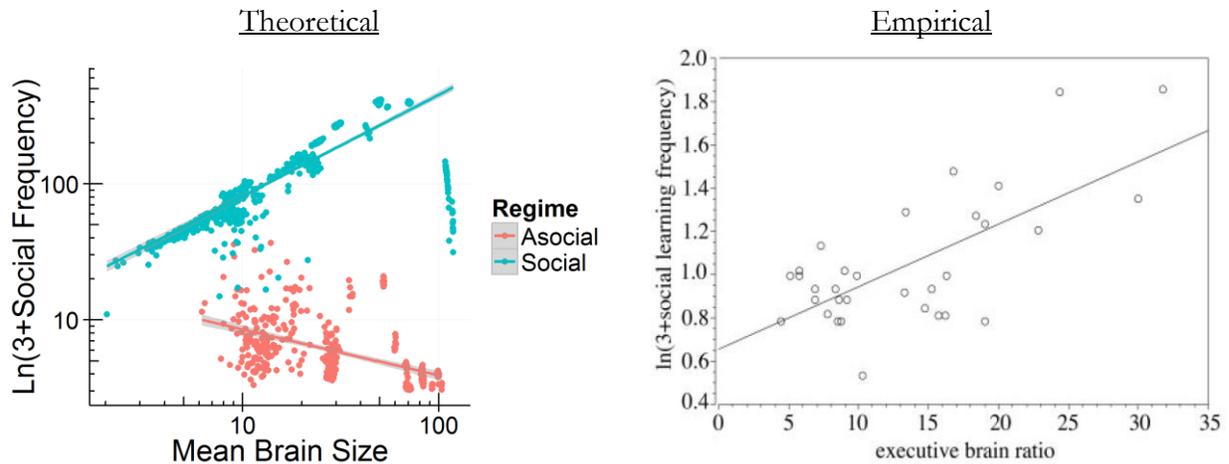


Figure 2.5. Example of correlations between brain size and incidences of social learning for regimes that emerge in our model (left). Correlation between brain size and incidences of social learning in the empirical literature (right; Reader & Laland, 2002).

Since the Cultural Brain Hypothesis predicts that brains are primarily for storing and managing information and that social learning is a cheaper way to acquire that information, then we should expect to see a correlation between brain size and social learning frequency. Moreover, we should expect to see intercorrelations between brain size and other features that support social learning, such as the length of the juvenile period, group size, and the amount of adaptive knowledge. Some of these have been measured empirically and we test these in the next three sections.

2.3.2.3 Brain Size and Juvenile Period

The Cultural Brain Hypothesis does not explicitly model the length of the juvenile period, but does include 2 periods of learning. In the first period, individuals can learn socially from their genetic parent or asocially by themselves. In the second period, individuals with a low s value are likely to update their knowledge asocially, while those with a higher s values only updated their

knowledge obliquely based on their v value - individuals had a $1 - s$ probability of updating asocially, sv probability of updating socially and an $s - sv$ probability of doing no further learning. Thus, sv represents a juvenile period in which learners could use payoff biased oblique transmission to update their knowledge. Positive relationships between brain size and the length of the juvenile period (weaning age to sexual reproduction) have been shown directly in primates (Charvet & Finlay, 2012; Joffe, 1997; Walker et al., 2006) and indirectly via age to sexual maturity in a variety of taxa (Isler & van Schaik, 2009). Empirical studies show a correlation of $r = .61, p = .037$ for primates (Joffe, 1997).

Our model indicates that among species that primarily rely on social learning, the relationship between brain size and the length of the juvenile period is moderate, ranging from $r = .35$ to $r = .59$. Our model predicts that among species that primarily rely on asocial learning, the relationship between brain size and an extended juvenile period is negative for the same reason as the negative relationship with social learning. But of course the Cultural Brain Hypothesis would suggest that such an extended juvenile period would be effectively non-existent among truly asocial learning species (as shown in Figure 2.6) – an extended juvenile period evolves for more opportunities to engage in social learning.

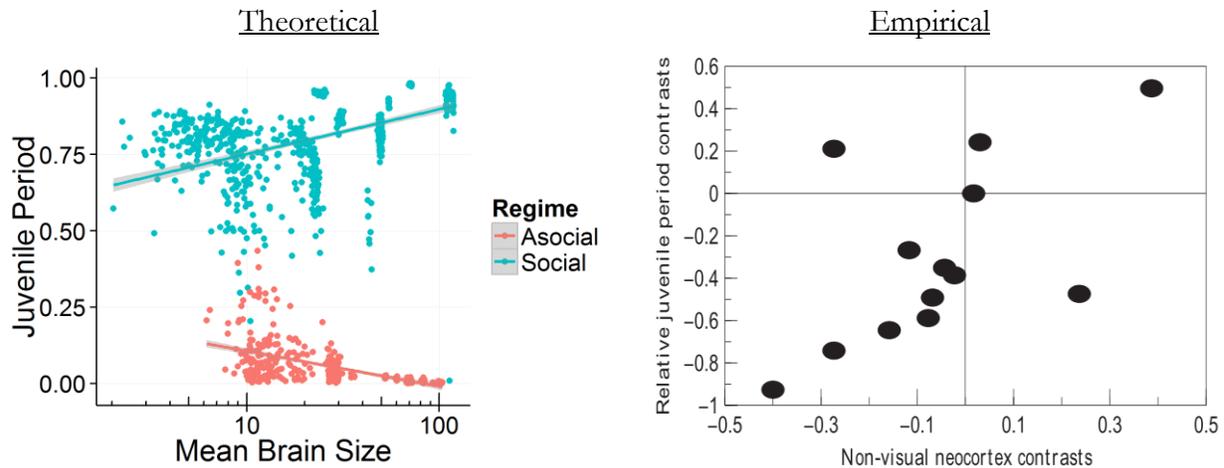


Figure 2.6. Example of correlations between brain size and the length of the juvenile period for regimes that emerge in our model (left). Correlation between brain size and juvenile period in the empirical literature (right; Joffe, 1997).

Joffe (1997) also measure the relationship between group size and the length of the juvenile period, offering another test of the Cultural Brain Hypothesis predictions.

2.3.2.4 Group Size and Juvenile Period

The Cultural Brain Hypothesis predicts a positive relationship between group size and the length of the juvenile period. This positive relationship is an indirect consequence of social learners having access to more knowledge in larger groups, creating a greater selection pressure for a longer juvenile period in which to take advantage of this knowledge. This in turn raises the average adaptive knowledge of the group, allowing for larger groups.

Joffe (1997) report a positive relationship between absolute juvenile period length and mean group size of $r = .57, p = .007$. A relative measure (like that used for the correlation between brain size and the length of the juvenile period) would have been more appropriate, since there is likely a tradeoff between primarily learning (during the juvenile phase) and exploiting knowledge (during the adult phase), but this too is likely positive (though perhaps smaller). Our model indicates that among

species that primarily rely on social learning, the relationship between group size and the juvenile period ranges from $r = .23$ to $r = .50$, but the relationship is effectively non-existent among those species who primarily rely on asocial learning. According to our theory, group size is unaffected by the juvenile period in these species since group size primarily supports the acquisition of more or better adaptive knowledge via social learning. We plot these relationships for the parameters chosen in Table 2.1 and 2.2 in Figure 2.7 below. Joffe (1997) do not provide a comparison plot, but we have generated one from her data.

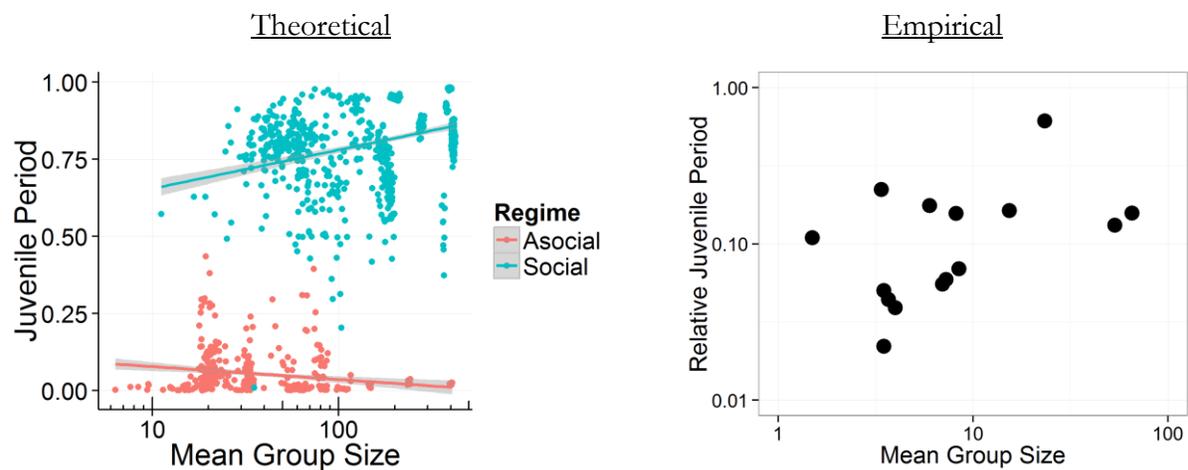


Figure 2.7. Example of correlations between group size and the length of the juvenile period for regimes that emerge in our model (left). Correlation between mean group size and relative juvenile period in the empirical literature (right; Joffe, 1997).

According to the Cultural Brain Hypothesis, the relationship between group size and the length of the juvenile period is indirect. The juvenile period is extended to allow for the acquisition of more adaptive knowledge via social learning and larger groups provide access to more models from which individuals can socially learn. The relationship between group size and adaptive knowledge is bidirectional. Groups with access to more adaptive knowledge can access more

calories, better avoid predators, and so on and are therefore larger. These larger groups give potential oblique learning social learners access to more models from which to learn creating more adaptive knowledge. Thus the last relationship we are able to test based on the existing empirical literature, the relationship between group size and adaptive knowledge, is unsurprising, but we include it for completeness.

2.3.2.5 Group Size and Adaptive Knowledge

The Cultural Brain Hypothesis predicts a positive relationship between adaptive knowledge and group size for all species. Species better able to evade predators, access calories, and avoid illness should be able to sustain larger groups. However, the relationship between group size and adaptive knowledge should be larger for species with access to social learning since these species can also exploit the greater amount of adaptive knowledge available in larger groups.

The relationship between adaptive knowledge, measured for example by toolkit size, and sociality has not been explored across species, but it has been examined within humans, chimpanzees, and orangutans (Lind & Lindenfors, 2010; van Schaik et al., 2003; Whiten & Van Schaik, 2007). Our model reveals a consistently positive relationship in both regimes between $r = .56$ and $r = .93$. We do not have comparisons for all these predictions nor comparisons across species, but empirical studies show a correlation of $r = .81, p = .005$ between group size and toolkit size for humans (Kline & Boyd, 2010) and a Spearman's correlation of $r_s = .87, p = .010$ between female group size and number of cultural traits in chimpanzees (Lind & Lindenfors, 2010). Most other chimpanzee and orangutan research tend to use association time as a measure of sociality, rather than group size, and thus are more difficult to directly compare. We plot these relationships for the parameters chosen in Table 2.1 and 2.2 in Figure 2.8 below.

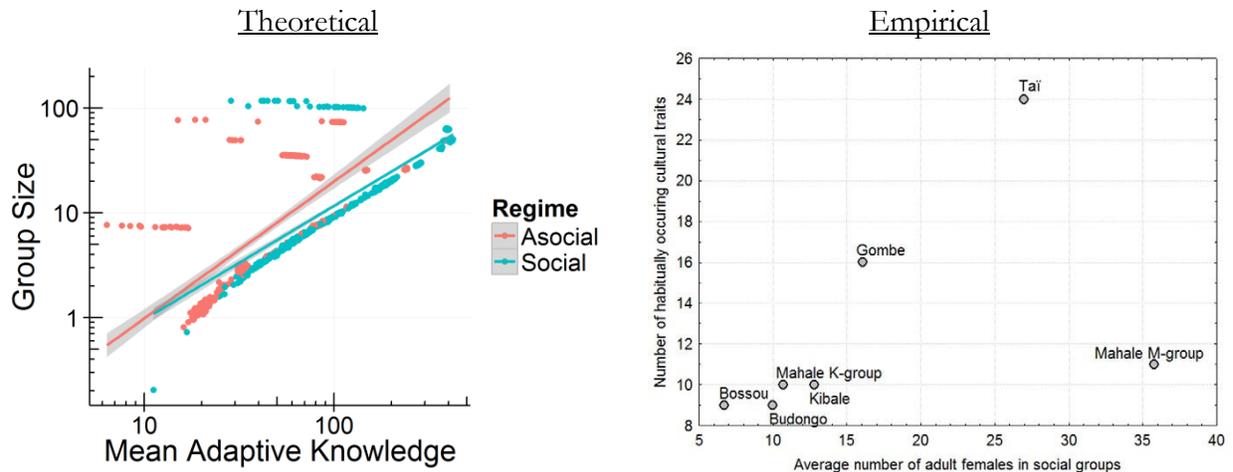


Figure 2.8. Example of correlations between group size and the mean adaptive knowledge for regimes that emerge in our model (left). Female group size and number of cultural traits in the empirical literature (right; Lind & Lindenfors, 2010).

2.3.2.6 Summary

In summary, the existing literature supports the predictions of the Cultural Brain Hypothesis. The Cultural Brain Hypothesis is a more general theory for a variety of observed empirical relationships. In the next section we explain in detail why we should expect these relationships (and others) to exist.

2.3.3 Processes Underlying Empirical Relationships

The causal relationships predicted by the Cultural Brain Hypothesis are outlined in Figure 2.9 below. Larger brains allow for more adaptive knowledge. More adaptive knowledge can in turn exert a selection pressure for larger brains. More adaptive knowledge allows for larger groups and creates a selection pressure for social learning to take advantage of the adaptive knowledge. Larger groups have more adaptive knowledge exploitable by those with social learning abilities. Ergo large groups of individuals who primarily rely on social learning have larger bodies of knowledge, exerting a selection pressure for an extended juvenile period in which more adaptive knowledge can be learned

(and created). An extended juvenile period creates a selection pressure for more reliance on oblique learning (learning from non-genetic parents in the group), which in turn creates a selection pressure for learning biases better able to select individuals and knowledge to learn. Oblique learning and learning biases lead to the realm of cumulative cultural evolution. These and other prerequisites are discussed in Section 2.4. Thus the Cultural Brain Hypothesis predicts that these measures are all positively intercorrelated among taxa with some amount of social learning, but are generally weaker or non-existent among taxa with little social learning. There has been less empirical data published for species with little social learning, perhaps due to a bias toward publishing statistically significant relationships.

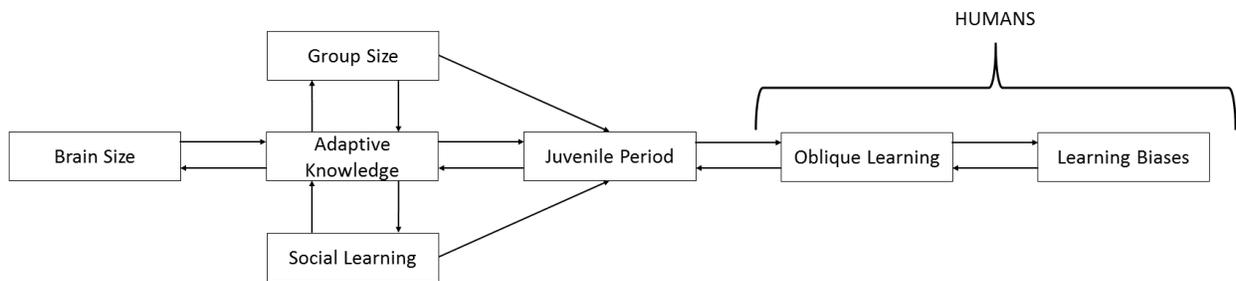


Figure 2.9. Causal relationships suggested by Cultural Brain Hypothesis.

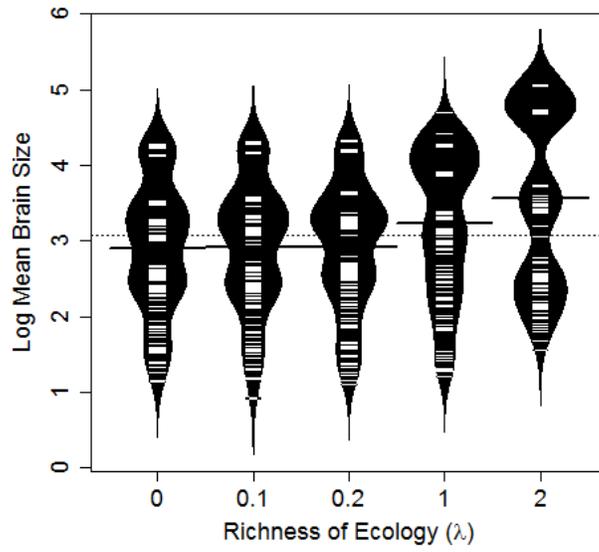
The evolution of different regimes is dependent on other factors modeled as parameters in our model. These include ecological factors such as the richness of the ecology (λ) as well as other factors that are themselves products of evolution. These include reproductive skew or mating structure (φ), migration rate (m), transmission fidelity (τ) and asocial learning efficacy (ζ). Other models have theorized the evolution of these structures, tendencies, and abilities, but here we are interested in the effect of these factors on the co-evolutionary processes shown in Figure 2.9. Under some conditions, these parameter values lead to the realm of cumulative evolution and this is the basis for the Cumulative Cultural Brain Hypothesis.

2.3.3.1 Richness of the Ecology (λ)

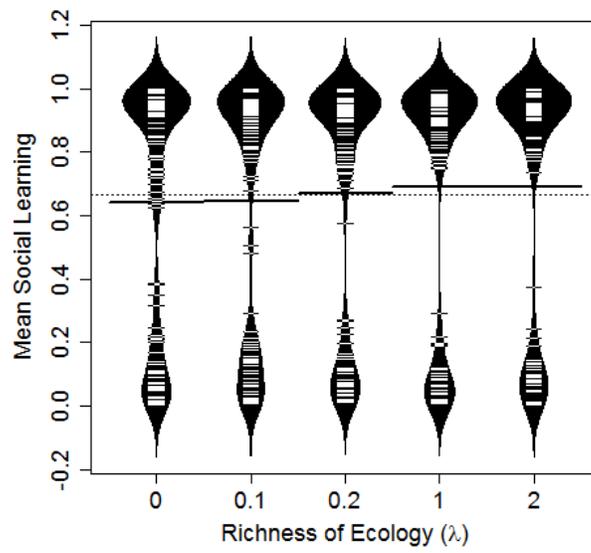
The richness of the ecology (λ) affects the range of possible brain sizes by adjusting the “bang for buck” on adaptive knowledge. That is, with higher values of λ less adaptive knowledge is needed to unlock more calories, evade more predators, and so on, allowing for larger brains, as shown in Figure 2.10(a) below. In doing so, it increases the strength of viability selection. This increased strength of selection leads to a slightly higher rate of social learning, shown in Figure 2.10(b), but also higher extinction rate, shown in Figure 2.10(c). In Figure 2.10(a) we see 3 regimes for brain size emerge. The regime with the largest brains corresponds to the realm of cumulative cultural evolution, the evolution of which is explained by the Cumulative Cultural Brain Hypothesis.

Encephalization slopes vary across mammalian taxa. Shultz and Dunbar (2010a) show that these slopes are predicted by sociality. The Cultural Brain Hypothesis suggests that the richness of the ecology⁸ may be one factor that predicts both this encephalization slope and sociality. Since we model the entire evolutionary process, we can plot these encephalization slopes and see the differential effect of λ on brain size and social learning (Figure 2.11). Higher λ has stronger selection and therefore more asocial learning when asocial learning is adaptive (little adaptive knowledge in group and more social learning when social learning is valuable (more social learning in the group)).

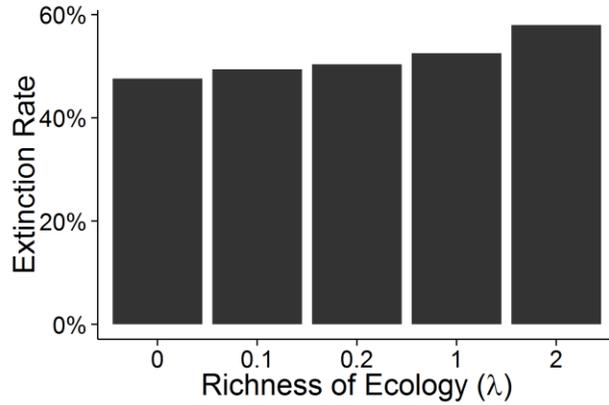
⁸ Richness of ecology is a mostly exogenous factor, but of course ecology (as opposed to environment) is an interaction between an organism and the environment (a spider and you in the same room have the same environment, but a very different ecology). Moreover, the organism and the environment can co-evolve (see Odling-Smee, Laland, & Feldman, 2003).



(a)

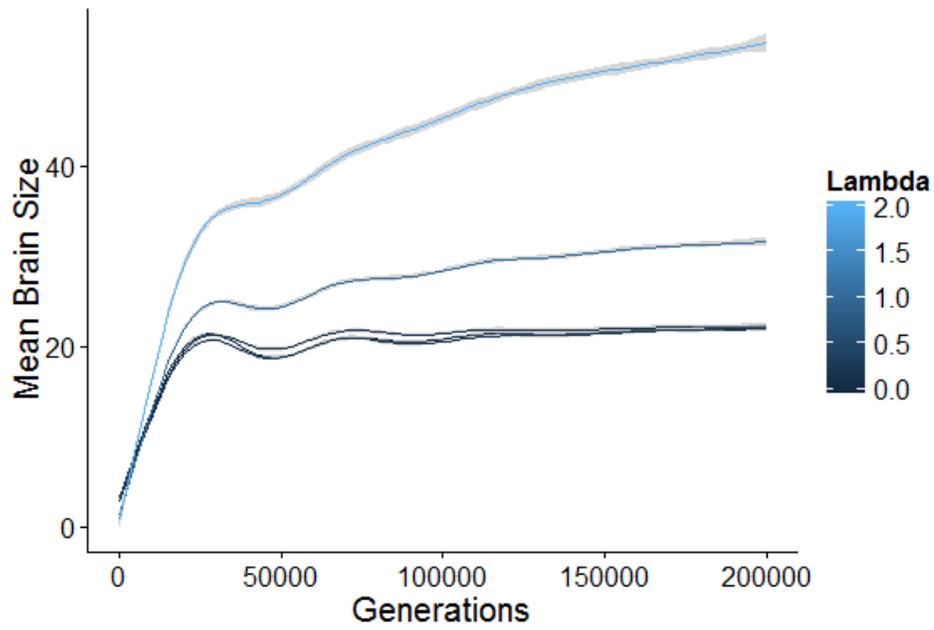


(b)

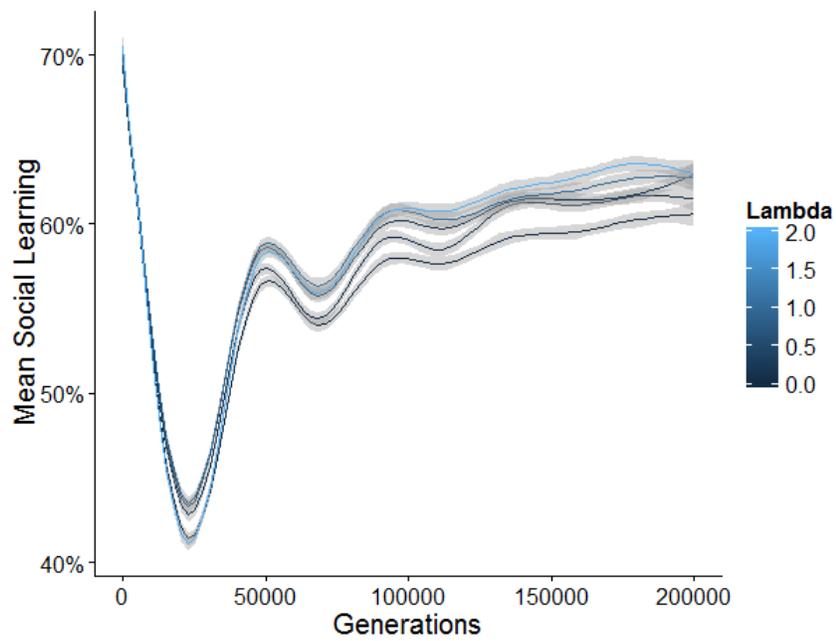


(c)

Figure 2.10 Bean plots showing the distribution of (a) brain size and (b) social learning means for different values of λ . The dotted horizontal line shows the global mean and the bolded horizontal lines show the group means. Bean plots show the distribution of values. We see 3 regimes emerge for brain size. The regime with the largest brain size corresponds to the realm of cumulative cultural evolution (see Section 2.4). (c) Plot showing the rate of extinction for different values of λ . Note that in all cases, $\lambda = 2$ has not yet reached equilibrium.



(a)



(b)

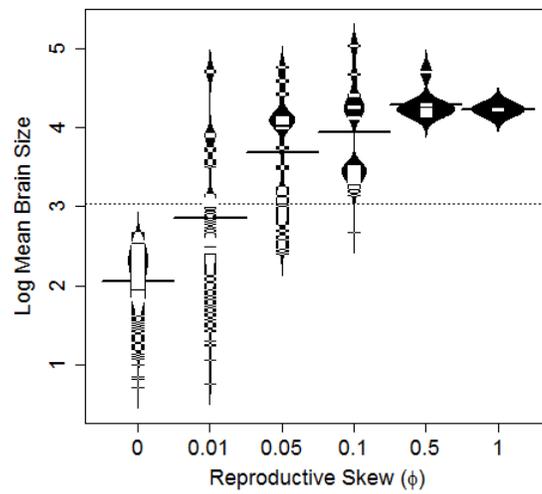
Figure 2.11 (a) Mean brain size showing the encephalization slope for different values of λ . Higher values have a larger encephalization slope. Note that brain size initially shrinks before growing

again. This shrinkage corresponds to the transition from asocial to social learning, resulting in more efficient brains – that is smaller brains are needed to acquire more adaptive knowledge if that knowledge is acquired socially. There is some evidence that the human brain has been shrinking in the last 10,000 to 20,000 years, which may be evidence that our species is not at equilibrium (Henneberg, 1988). (b) Mean social learning over generations. There are two features of note. First, we started our simulations with social learning to show that social learning is maladaptive in the absence of adaptive knowledge. Asocial learners quickly invade. It is only when asocial learners have generated sufficient adaptive knowledge that social learners again have an advantage. Since we know that two regimes reliably emerge, mean social learning in these plots represents the relative number of conditions in which social and asocial learners emerge rather than a value of social learning characteristic of the world.

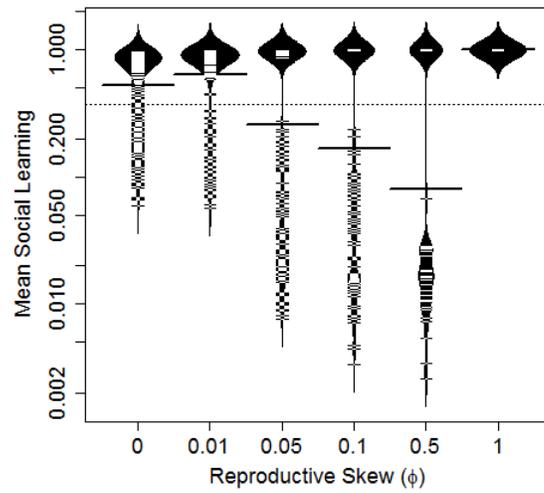
2.3.3.2 Reproductive Skew or Mating Structure (φ)

Brain size correlates with mating structure in mammalian and avian lineages (Shultz & Dunbar, 2010a; Shultz & Dunbar, 2010b). We model the effect of mating structure or reproductive skew using φ . The φ parameter affects the relationship between individual adaptive knowledge and the mating competition. When $\varphi = 0$, all individuals have the same probability of reproducing regardless of their adaptive knowledge. This corresponds to a perfectly monogamous society. As φ increases, we enter into a slightly “monogamish” or cooperative breeding society and then to a polygynous society for very high values of φ . In doing so, it increases the strength of selection for more adaptive knowledge, but the results of this increase in fecundity selection may be surprising. First, brain size increases with φ (Figure 2.12a), but this relationship is misleading, because the extinction rate also increases with higher φ (Figure 2.12c). Extinction rates go up, because variance is reduced with too high fecundity selection. More adaptive knowledge is sought at any cost, but in a world with little adaptive knowledge, the best way to acquire this knowledge is via asocial learning.

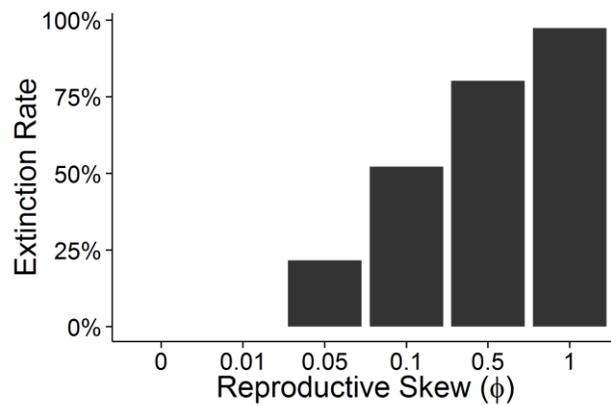
This leads to populations getting stuck in the world of asocial learning without the necessary variance (some social learning) to take advantage of the existing body of adaptive knowledge. Second, Figure 2.12(b) reveals for these same reasons, the tendency to use social learning decreases with greater reproductive skew. The highest amount of social learning occurs in monogamish or cooperative breeding populations. We return to this in Section 2.4 on the Cumulative Cultural Brain Hypothesis.



(a)



(b)

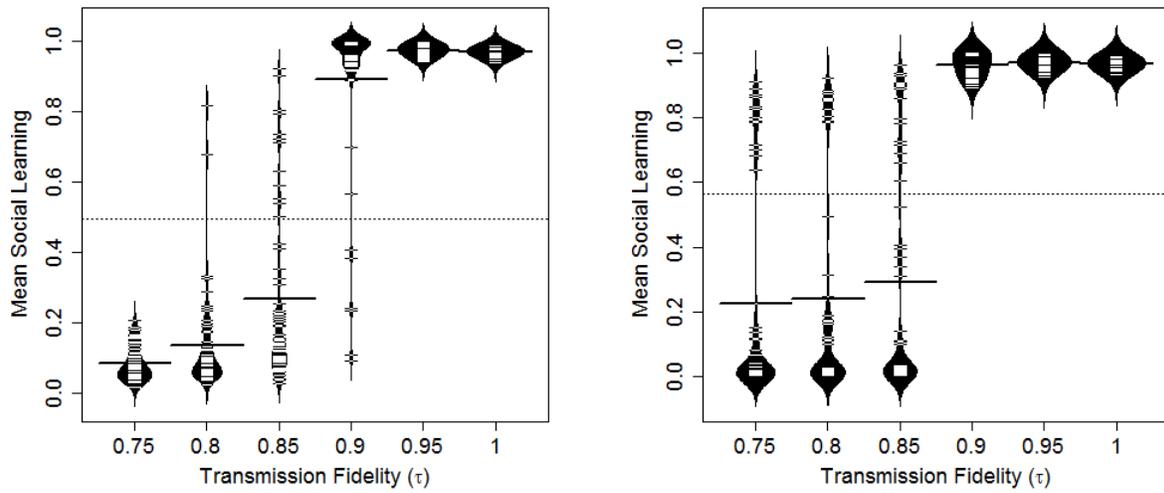


(c)

Figure 2.12 Bean plots showing the distribution of (a) brain size and (b) social learning means for different values of ϕ . The dotted horizontal line shows the global mean and the bolded horizontal lines show the group means. Bean plots show the distribution of values. W (c) Plot showing the rate of extinction for different values of ϕ .

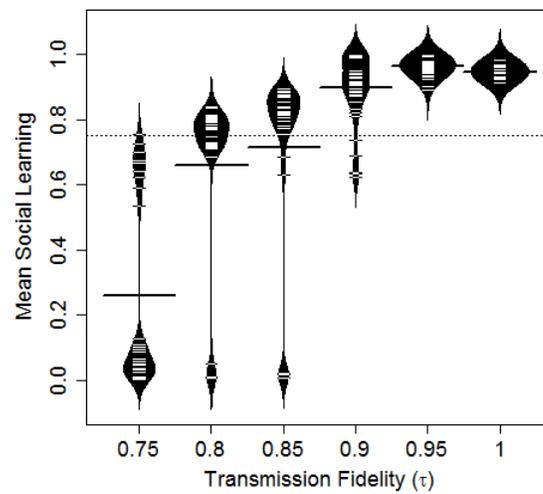
2.3.3.3 Transmission Fidelity and Asocial Learning Efficacy (τ and ζ)

Transmission fidelity (τ) affects how much loss takes place in the transmission of adaptive knowledge from genetic parent or cultural model (if oblique learning evolves) to an individual in the next generation. Asocial learning efficacy (ζ) affects the efficiency with which individuals can generate new adaptive knowledge based on their own brain size. These two parameters interact in interesting ways to affect the evolution of social learning with consequent effects on brain size, population size, etc. As Figure 2.13 shows, unsurprisingly, higher transmission fidelity leads to more social learning. However, we see an interaction, where social learning is more likely to evolve when asocial learning efficacy is higher. These simulations exaggerate this overall pattern, because we begin with social learners, but the key message is that social learners stand on the shoulders of effective asocial learners whose knowledge they exploit. We will return to this in the next section on the Cumulative Cultural Brain Hypothesis.



$$\zeta = 0.1$$

$$\zeta = 0.4$$



$$\zeta = 0.7$$

Figure 2.13 Bean plots showing the distribution of social learning for different values of τ and ζ . The dotted horizontal line shows the global mean and the bolded horizontal lines show the group means. Bean plots show the distribution of values. Unsurprisingly, higher transmission fidelity leads to more social learning. However, we see an interaction, where social learning is more likely to evolve when asocial learning efficacy is higher. These simulations exaggerate this overall pattern, because

we begin with social learners, but the key message is that social learners stand on the shoulders of effective asocial learners whose knowledge they exploit.

2.4 Cumulative Cultural Brain Hypothesis

Beyond the hypothesis that social learning, brain size, adaptive knowledge, and group size may have coevolved so as to create the patterns found in the empirical literature, we are also interested in the conditions under which these variables might interact synergistically to create highly social species with large brains and substantial accumulations of adaptive knowledge (humans). To assess when an accumulation of adaptive knowledge becomes cumulative cultural evolution, we apply a standard definition of cumulative cultural products, as being those products that a single individual could not invent by themselves in their lifetimes. To calculate this for our simulated species, we ask what the probability is that an individual with the average brain size of the species would invent the mean level of adaptive knowledge in that species via asocial learning. Formally, this is given by (8).

$$\int_a^{\infty} N(\zeta b_{ij}, \sigma_a \zeta b_{ij}) \quad (8)$$

We can set an arbitrary threshold, for example, where the probability of any individual acquiring this level of adaptive knowledge through asocial learning is less than 0.1%. At this level, for an entire population to develop that level of adaptive knowledge through asocial learning is 0.001^{N_j} , i.e. exceedingly unlikely. Thus mean levels of adaptive knowledge that are so exceedingly unlikely to have been acquired through asocial learning can be attributed to cumulative cultural evolution. In Figure 2.14 below, we plot brain size against the probability of acquiring that amount of information.

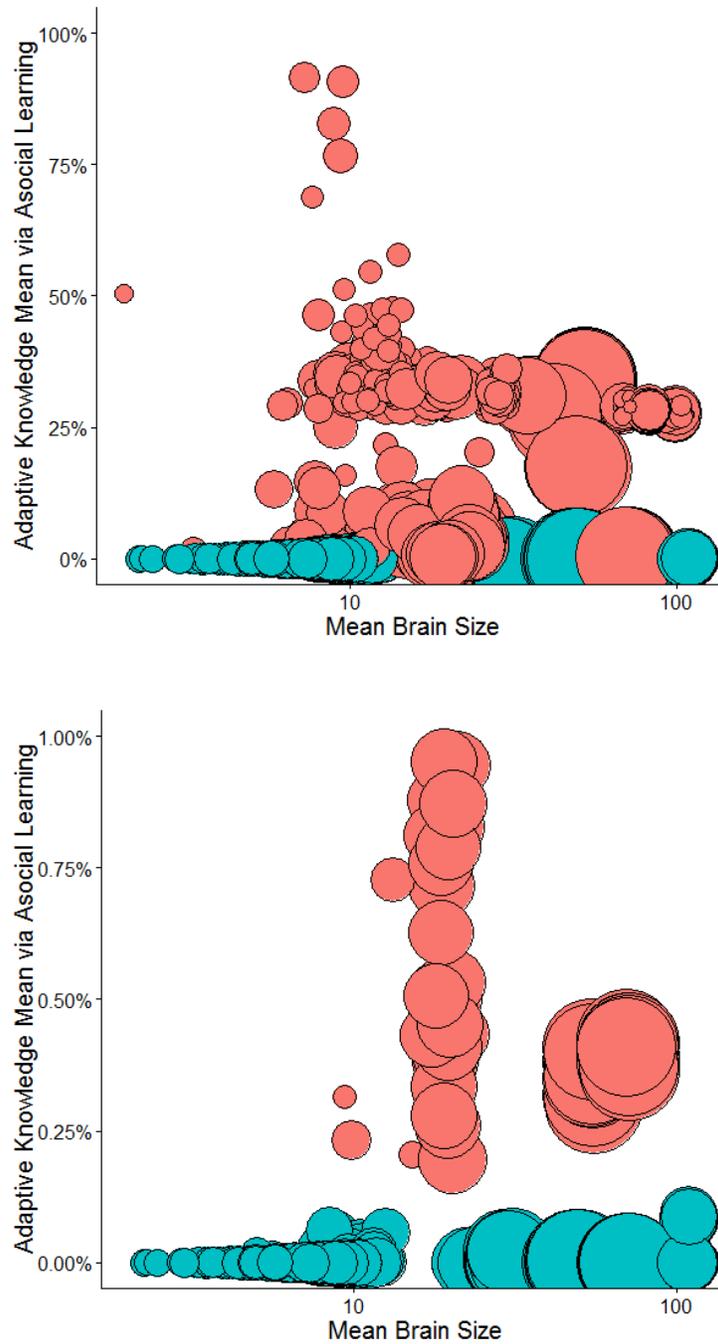


Figure 2.14 Log mean brain size against the probability of acquiring the mean adaptive knowledge in the group via asocial learning for $\lambda = 1$. Circle size indicates the mean population size. Red simulations did not cross the threshold into cumulative cultural evolution, but aquamarine simulations did. In the lower graph, we zoom into the 0% to 1% probability range.

Next, we can look at what parameters increase and decrease the probability of entering into the realm on the bottom right corner of Figure 2.14. Large brained species with a lot of adaptive knowledge, which they were unlikely to acquire without cumulative cultural evolution; the human species. Note that these species tend to have very large populations, for the reasons outlined in the previous section.

2.4.1 Transmission Fidelity Drives Larger Brains

The first prediction of the Cumulative Cultural Brain Hypothesis is that transmission fidelity is the key to entering into the realm of cumulative cultural evolution. Figure 2.15 plots the percentage of simulations with cumulative cultural evolution for different values of τ . We see a threshold effect, where for very high fidelity transmission ($\tau > 0.85$), social learning and large brains evolve under a wide range of parameters.

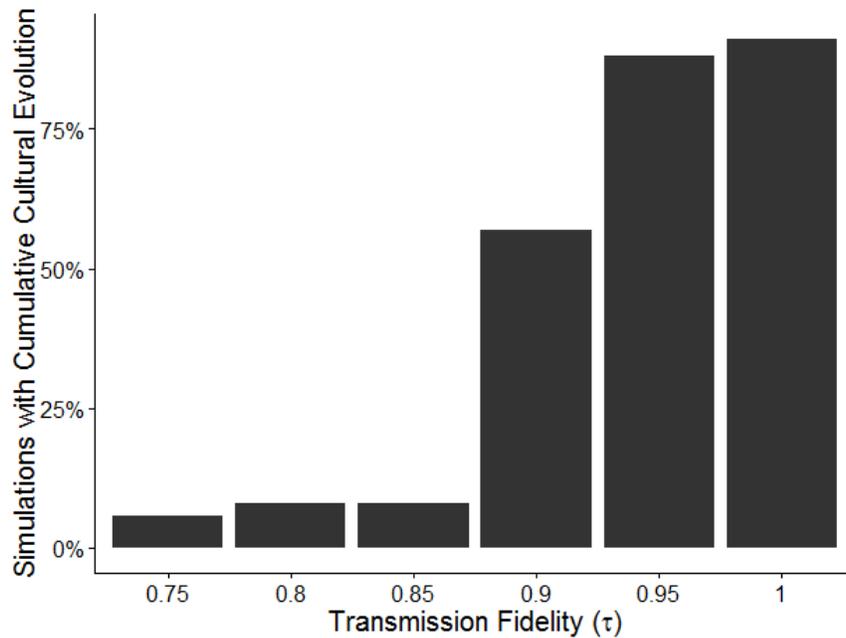


Figure 2.15 Percentage of simulation which enter the realm of cumulative cultural evolution.

Embedded in τ , and eventually oblique learning and learning bias, is cognitive abilities like theory of mind, the ability to recognize, distinguish, and imitate potential models, but also teaching and social tolerance. Cultural learning can also increase τ . For example, culture can increase social tolerance by expanding the circle of moral regard (members of the ingroup) or through cultural innovations like teaching (Kline, 2014; Morgan et al., 2015). Such cultural innovations may be a result of the inverse relationship between transmission fidelity and cultural complexity. As cultural complexity goes up, by definition, culture is more difficult to acquire and therefore transmission fidelity decreases. In our model we fixed τ for each simulation run, reducing the parameter space.

There are few empirical measures of transmission fidelity. Claidière and Sperber (2010) review several studies and based on these estimate human transmission fidelity at 86%. Humans are likely capable of at least 86% fidelity, but the fidelity depends on other factors, including difficulty of the task. Other animals, such as Norway rats, are capable of transmission of up to 95%. Norway rats don't possess cumulative cultural evolution, thus transmission fidelity is necessary, but not sufficient for large brained social learners with cumulative culture.

2.4.2 Social Learning Benefits from Smart Ancestors

We find an interaction between transmission fidelity τ and individual learning ζ . If ζ is too high, individual learning is too efficient and social learning struggles to take off, except at very high rates of transmission, but if ζ is too low, even if social learning out-competes individual learning, the population have smaller brains and less adaptive knowledge compared to when social learning out-competes more effective individual learning. These results suggest that social learners stand on the shoulders of effective asocial learners. That is, when social learning can initially exploit the adaptive knowledge developed by more effective individual learning, social learning results in larger brains.

To explore this further we ran our simulations starting with 100% social learners (see Figure 2.11b). Here social learning is initially beaten out by asocial learning since there is no adaptive knowledge in the world, but quickly social learning allows individuals to exploit this knowledge and begin the autocatalytic coevolution of brains, population size, and adaptive knowledge gained through social learning. The Cumulative Cultural Brain Hypothesis predicts highly innovative ancestors; the kind of individual innovativeness we see in chimpanzees.

2.4.3 Mating Structure Matters

As we discussed in Section 2.3.3.2, monogamish or cooperative mating structures are more likely to lead to social learning and therefore to cumulative cultural evolution. Too strong a selection pressure leads to bigger brains via asocial learning that often go extinct (even when we start with fully developed social learning!). We graph the probability of entering into the realm of cumulative cultural evolution for different values of φ in Figure 2.16. We see a Goldilock's zone around $\varphi = 0.01$. As reproductive skew (e.g. polygyny) increases, asocial learning is favored and entering the realm of cumulative cultural evolution is less likely. However, perfect monogamy is also less likely to favor social learning and entering the realm of cumulative cultural evolution than a monogamish or cooperative breeding social structure.

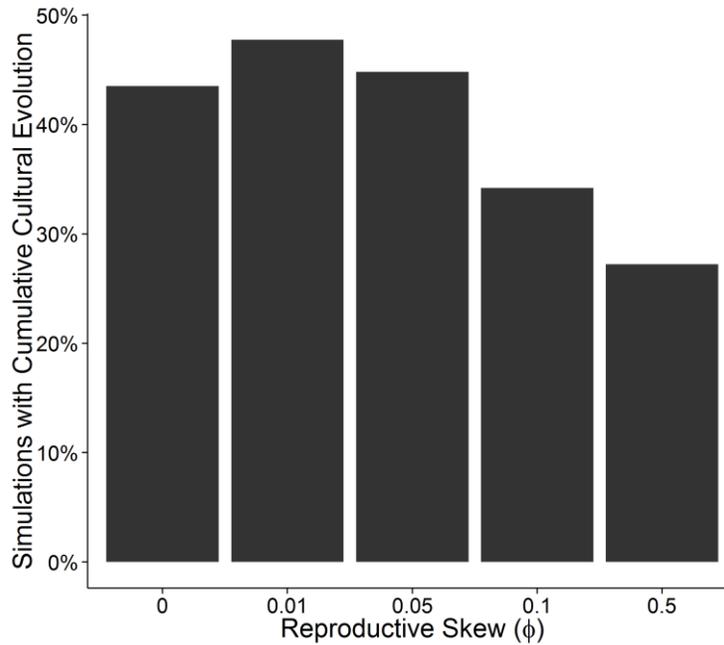


Figure 2.16 Percentage of simulations which enter the realm of cumulative cultural evolution.

2.4.4 Asocial Learning Can Lead to Larger Brains

Our model suggests two pathways to larger brains - via high fidelity social learning or very effective asocial learning. However, while the social learning pathway leads to an equilibrium brain size, the asocial learning pathway usually leads to extinction. The results of these are visible in the rise then fall of average brain size as social learning moves to fixation. Here asocial learning leads to larger brains, but these individuals are fundamentally limited in their ability to fill their brains by the effectiveness of asocial learning (ζ). The selection pressure for more adaptive knowledge eventually leads to larger and larger brains. Individuals with these very large brains are more likely to die due to an increasing death rate. In contrast, although social learning individuals have smaller brains initially compared to their individual learning brethren, they are able to more fully fill these brains through social learning - their brains are more efficient. If transmission fidelity is high enough (τ is a parameter in our model, but it is likely that the package of abilities allowing for high transmission

fidelity evolved at some point), social learning eventually allows for brain sizes comparable to those achieved by asocial learning. In contrast, if transmission fidelity is lower, brains are smaller than the large brains achieved by individual learning. This model suggests that social learning can evolve with smaller brains and out-compete larger brained individuals using individual learning and that transmission fidelity is the key to large human brains.

2.4.5 The Richness of the Environment Can Give You a Bigger Brain

Brains are costly, but this cost can be offset by more adaptive knowledge. The degree of mitigation is determined by λ . We find that higher λ values allow for the evolution of larger brains (see Figure 2.10). Basically, you need to be in an environment where adaptive knowledge pays off well enough to pay for those costly brains.

2.5 General Discussion

In this discussion section we (1) summarize our key findings, (2) review these findings in the context of the cultural intelligence hypothesis and other related work, (3) discuss limitations of this work and ongoing inquiries, and (4) consider the implications of our results.

2.5.1 Summary of Key Findings

Our model provides a potential evolutionary mechanism that can explain a variety of empirical findings linking brain size, group size, toolkit size, social learning, mating structures, and developmental trajectory as well as brain evolution differences among species. It can also explain the different encephalization slopes that have been found in different taxa and help explain why brain size correlates with group size in some taxa, but not others. The key message of the Cultural Brain Hypothesis in contrast to competing explanations is that brains are primarily for the storage and management of adaptive knowledge and that this adaptive knowledge can be acquired via asocial or social learning. Social learners flourish in an environment filled with knowledge (such as those found

in larger groups and those that descend from smarter ancestors) whereas asocial learners flourish in environments where knowledge is scarce. The correlations that have been found in the empirical literature between brain size, group size, social learning, the juvenile period, and adaptive knowledge, are an indirect result of these processes. The Cumulative Cultural Brain Hypothesis posits that these very same processes can, under very specific circumstances, lead to the realm of cumulative cultural evolution. These circumstances include when transmission fidelity is sufficiently high, reproductive skew is in a Goldilocks zone close to monogamy, effective asocial learning has already evolved, and the ecology offers sufficient rewards for adaptive knowledge. In making these predictions, the Cultural Brain Hypothesis and Cumulative Cultural Brain Hypothesis tie together several lines of empirical and theoretical research

2.5.2 The Model in the Context of Other Theory and Findings

Our model is consistent with relevant earlier models by Pradhan et al. (2012) and Gavrilets and Vose (2006). Pradhan et al. (2012) modeled mechanisms of increased technological complexity as mediators for the relationship between population and brain size. Our model focuses on the broader social factors underlying these mechanisms and thereby explains a wider range of phenomena. Gavrilets and Vose (2006) modeled an interpretation of the Machiavellian intelligence hypothesis. Their simulation is Machiavellian insofar as it models male-male competition for females, but rather than directly relying on more sophisticated cognitive mechanisms, males have an evolving learning ability and cranial capacity, which allows them to learn strategies, which are used to compete for females. These strategies are invented and forgotten by individuals, presumably through asocial learning. In formalizing the Machiavellian intelligence hypothesis in this way, Gavrilets and Vose (2006) model captures some of the elements expressed in some of the foundational papers on the Social Intelligence and Machiavellian Intelligence hypothesis (Alison, 1966; Humphrey, 1976;

Whiten & Byrne, 1988a; 1988b; for a more recent discussion, see Whiten & van Schaik, 2007). The model also captures some of the dynamics of the Cultural Brain Hypothesis – the relationship between learning ability, cranial capacity, and competitive strategies. However, their model ignores the mechanisms of cultural learning through which adaptive knowledge is created and transmitted. We focus on both in our model and show the importance of the way in which adaptive knowledge is socially transmitted.

The predictions made by the Cultural Brain Hypothesis are consistent with other theoretical work focused on parts of the theory. For example, researchers including Henrich (2004b), Powell, Shennan, and Thomas (2009), and Kobayashi and Aoki (2012) have argued for the causal relationship between sociality and the complexity and amount of adaptive knowledge. Similarly, the predictions made by the Cumulative Cultural Brain Hypothesis are also consistent with other theoretical work focused on parts of the theory. For example, researchers including Enquist, Strimling, Eriksson, Laland, and Sjostrand (2010) and Lewis and Laland (2012) have argued for the importance of high fidelity transmission for the rise of cumulative cultural evolution.

Cultural variation is common among many animals (e.g. rats, pigeons, chimpanzees, and octopuses), but cumulative cultural evolution is rare (Boyd & Richerson, 1996). Boyd and Richerson (1996) have argued that although learning mechanisms, such as local enhancement (sometimes classified as social learning) can maintain cultural variation, observational learning is required for cumulative cultural evolution. Our model supports this argument showing that only high fidelity observational social learning allows for cumulative cultural evolution. In our model, cumulative cultural evolution exerts a selection pressure for larger brains which in turn allows more culture to accumulate. Prior research (e.g. Dean et al., 2012; Heyes, 2012; Morgan et al., 2015) has identified

many mechanisms, such as teaching, imitation, and theory of mind, underlying high fidelity transmission and cumulative cultural evolution.

Our model reveals that in general, social learning leads to more adaptive knowledge and larger brain sizes, but shows that asocial learning can also lead to increased brain size. Further, our models indicate that asocial learning may provide a foundation for the evolution of larger brained social learners. These findings are consistent with Reader, Hager, and Laland (2011), who argue for a primate general intelligence that may be a precursor to cultural intelligence and correlates with absolute forebrain volume.

In a recent empirical study of 36 species across many taxa, MacLean et al. (2014) show that brain size correlates with an ability to monitor food locations when the food was moved by experimenter and to avoid a transparent barrier using previously acquired knowledge in order to acquire food. They also show that brain size predicts dietary breadth, which was also an independent predictor of performance on these tasks. Brain size did not predict group size across all these species (some of whom relied more on asocial than social learning). These results are precisely what one would expect based on the Cultural Brain Hypothesis; brains have primarily evolved for the storage and management of adaptive knowledge. Moreover, the Cultural Brain Hypothesis predicts a strong relationship between brain size and group size among the social learning species, but a weaker or non-existent relationship among species that relied more on asocial learning. We plan to conduct this analysis in a future study.

Our simulation results are consistent with several lines of empirical data for brain size and group size among extant primates, but suggest a different mechanism for humans. Rather than a selection pressure for Machiavellian intelligence or tracking other individuals within the group, the extraordinarily large human brain may have evolved initially for this (theory of mind and other such

skills may be a prerequisite for social learning) and for asocial learning (similar to chimpanzee learning), but then entered an alternative evolutionary pathway. On this alternative human evolutionary pathway, the need to socially acquire, store, and organize adaptive knowledge resulted in the coevolution of brains and adaptive knowledge – a Cumulative Cultural Brain Hypothesis.

2.5.2.1 Cultural Intelligence Hypothesis

The Cultural Intelligence Hypothesis, deserves special mention. The Cultural Brain Hypothesis and Cumulative Cultural Brain Hypothesis are an attempt to formalize many of the ideas suggested by the closely related Cultural Intelligence Hypothesis. There is some amount of disagreement in the literature about the use of the label “Cultural Intelligence Hypothesis”. Whiten and Van Schaik (2007), who have temporal precedence, have used the “Cultural Intelligence Hypothesis” to argue that culture may have driven encephalization in non-human great apes with empirical evidence supporting this hypothesis. These data are consistent with the Cultural Brain Hypothesis. Herrmann et al. (2007) have used “Cultural Intelligence Hypothesis” to argue that humans have a suite of cognitive abilities that have allowed for the acquisition of culture, similar to a hypothesis previously referred to as the Vygotskian Intelligence Hypothesis (Moll & Tomasello, 2007). These data are consistent with the Cumulative Cultural Brain Hypothesis.

2.5.3 Limitations and Future Directions

There were several assumptions that simplified our model and made it less computationally costly. Future models may address some of these shortcomings and explore additional parameters. One such improvement is to explicitly track different cultural traits with different cognitive costs and fitness payoffs. By doing this, we could better explore the benefits to migration and cultural recombination. We would also like to more fully explore the impact of the relationship between adaptive knowledge and carrying capacity. Currently, the richness of the ecology only affects

individual survival based on paying the calorie cost of costly brains, but the richness of the ecology also affects the carrying capacity of the population with consequent effects for the dynamics between brain size, adaptive knowledge and population size.

Another future improvement we have previously mentioned is the endogenization of transmission fidelity (τ) and reproductive skew (φ). As previously discussed, these parameters are themselves subject to genetic and cultural evolutionary processes and thus ought to be modeled as endogenous variables. In our model, we can discuss the effect of different evolutionary outcomes or values of transmission fidelity and reproductive skew, but not their evolution.

Two or three regimes emerged in our models. Future models should explore the adaptive dynamics to determine the invasion fitness of the different equilibrium states discovered in our model. These models will help us better understand the evolutionary dynamics that may have occurred when different previously geographically separated hominin species encountered each other (e.g. the European encounter between modern humans and their larger brained Neanderthal cousins).

Finally, one result that deserves more exploration is the brain shrinkage that occurs during the transition from reliance on asocial learning to reliance on social learning. These results hint that the process underlying the Cultural Brain Hypothesis and Cumulative Cultural Brain Hypothesis may also help explain evidence suggesting that human brains have been shrinking in the last 10,000 to 20,000 years (Henneberg, 1988), which may be evidence that our species is not at equilibrium.

Chapter 3: Sociality Influences Cultural Complexity

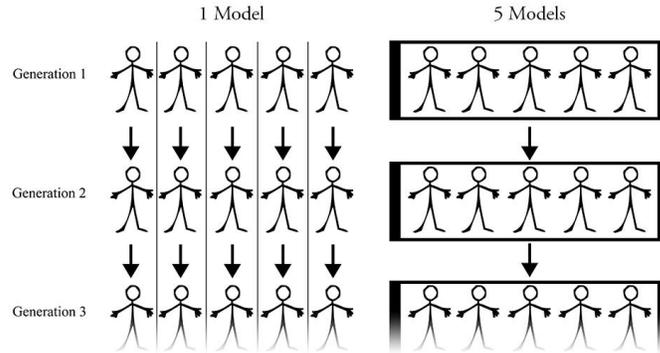
Humans may be unique among species in generating the cumulative cultural evolutionary processes that give rise to complex behavioural skills and technologies (Dean et al., 2012; Pagel, 2012; Pradhan et al., 2012; Whiten, Hinde, Laland, & Stringer, 2011). A growing class of theoretical models suggest that the emergence of such complex and ‘difficult to learn’ cultural traits (tools, techniques, and skills), such as many of the technologies used by hunter-gatherers, is heavily influenced by the abilities of learners to access a larger social network of other individuals (Aoki, Lehmann, & Feldman, 2011; Enquist et al., 2010; Henrich, 2004b; Hill et al., 2011; Kobayashi & Aoki, 2012; Lehmann, Aoki, & Feldman, 2011; Powell et al., 2009; Premo & Kuhn, 2010; Vaesen, 2012; van Schaik & Pradhan, 2003). On the empirical side, field evidence consistent with these models has begun to emerge. This evidence includes analyses of the complexities of toolkits among populations (Collard, Ruttle, Buchanan, & O’Brien, 2012; Kline & Boyd, 2010; van Schaik et al., 2003) as well as detailed studies of particular archaeological, ethnographic, and ethnohistorical cases (Edinburgh, 2009; Henrich, 2004b; Marquet et al., 2012; Powell et al., 2009; Wadley et al., 2011). Thus, technological sophistication may depend on sociality, on the size and interconnectedness of populations. This has led some to suggest that the key differences between human ancestors and other primates may lie in the domain of sociality and population or network structures (Henrich & McElreath, 2003; Herrmann et al., 2007; van Schaik & Burkart, 2011). Of course, there is every reason to suspect that other factors also influence cumulative cultural evolution in substantial ways (Collard, Buchanan, Morin, & Costopoulos, 2011; Collard, Kemery, & Banks, 2005; Nielsen, 2012).

Here, we test the relationship between sociality and cumulative cultural evolution in two laboratory experiments, where sociality is operationalized in terms of a participant’s ability to access and learn from multiple experienced individuals (‘models’ or ‘cultural parents’). Experiment 1 tests

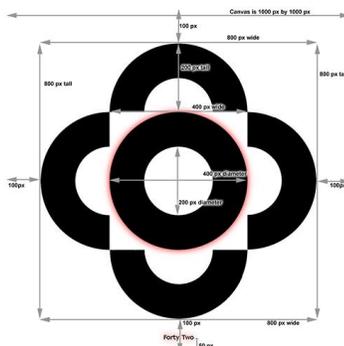
the effect of the number of accessible models on cumulative cultural change over successive laboratory generations using a first generation of untrained or ‘uncultured’ participants. Experiment 2 tests the effect of the number of models on the loss of cultural complexity over successive generations by beginning with a first generation of trained ‘highly cultured’ participants.

3.1 Methods

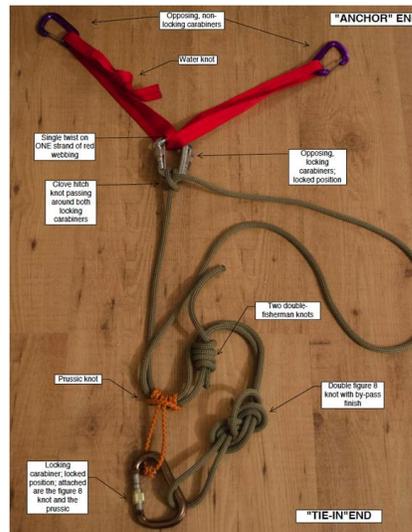
In both studies, we tested the transmission of knowledge and skill using undergraduates ($N = 100$) randomly assigned to one of two treatments (1-Model vs. 5-Models) each with 10 generations (5 participants per treatment per generation). In the 1-Model treatment, participants in generations 2 to 10 had access to information from only one participant from the previous generation (1 ‘cultural parent’). In the 5-Model treatment, participants in generations 2 to 10 had access to information from all five participants in the previous generation (5 ‘cultural parents’). Figure 3.1a illustrates our experimental design. Participants’ performance was incentivized with additional entries into a \$100 raffle when (1) they performed relatively better in their own generation and (2) those they transmitted to emerged as the best performer in the next generation. Thus, participants made the most money when both they and their cultural offspring performed the best in their respective generations. Below, we first briefly present the methods for each experiment, and then move onto the results.



(a)



(b)



(c)

Figure 3.1 (a) An illustration of the experimental design. (b) The target image for Experiment 1. Note the words “Forty Two” at the base of the image and the red glow around these words and the circle. Participants were not required to recreate the dimension arrows. (c) The knots used in Experiment 2. Participants were asked to tie this setup to two chairs. Larger versions of (b) and (c) can be found in Appendix A, Figures A.1 and A.2.

In Experiment 1, participants with little or no prior experience in image editing were asked to recreate a target image using a complex editing program called GIMP (2012). We also supplied a

second version of the target image with annotated measurements (as shown in Figure 3.1b). In all generations, participants were given sufficient time (up to 15 minutes) during which they were permitted to write up to two pages of information to assist the next generation. All generations, except Generation 1, were provided with the written information, the target image (with and without measurements) and a screenshot from their cultural parent or parents and given up to 25 minutes to recreate the target image. Those in the 1-Model treatment had access to only one participant's information and image while those in the 5-Model treatment had access to all five participant's information and images. Participants' ($N = 100$, 71 female) ages ranged from 15 to 35 ($M = 20.52$, $SD = 2.80$). Additional participant information is provided in Appendix A.1.

Each participant's final image was rated in two ways. First, each image was assessed by one of three human raters using a scale designed to measure the level of reproduction of various features of the image model (alignment, size, shape, gradient, etc.). Scores on our scale ranged from 0 to 59, which we rescaled to a percentage from 0 to 100. Inter-rater reliability, calculated on a range of images from a pilot study and images from participants who exceed our maximum experience threshold, was very high ($ICC(3, 1) = 0.997$). Appendix A.5 and A.6 provides information on the training and evaluation of raters.

Second, as a check on these human-ratings, final images were also assessed using a similarity algorithm (35, 36; see Appendix A.4 for details). The algorithm computes the normalized cross correlation metric, which yields a value between 0 and 1 for the two images by pairing them pixel by pixel and calculating a correlation. Ratings from this algorithm and our human raters were highly correlated ($r = 0.87$, $p < 0.001$). However, because the algorithm does not assess features clearly relevant to human minds (e.g. the target's degree of red glow, misalignment of image, etc.), we ran

our analyses below using the above human rating scale, and rely on the algorithm's overall similarity measure only as a robustness check.

In Experiment 2, participants were asked to tie a system of connected knots commonly used in rock climbing. Generation 1 was trained by the experimenter using standardized instructions to become 'experts' at tying this system of knots. Participants in all generations were given sufficient time (up to 20 minutes) during which they were permitted to create an instructional video detailing the tying and placement of each knot. To reduce experimenter bias, a camera was strapped to each participant's forehead, providing a first person view. All subsequent generations were provided with the video from the previous generation as well as the participant's score and were given up to 50 minutes to learn and recreate the knot system. The 1-Model treatment had access to only one participant's video and score while the 5-Model treatment had access to all five participant's videos and scores. Participants' ($N = 100$, 71 female) ages ranged from 17 to 37 ($M = 20.48$, $SD = 3.15$; further details in Appendix A.1).

To assess the performance of each participant, their final knot system was assessed by one of two human raters, using a custom rating scale inspired by a scale used to assess sutures when training surgeons (Tytherleigh, Bhatti, Watkins, & Wilkins, 2001). The scale was used to assess the deviation of each knot and knot position from the original model. The scale scores ranged from 0 to 37, which we rescaled to percentages (see Appendix A.6 for the complete scale). Inter-coder reliability, calculated on a range of knots from a pilot study, was very high ($\alpha = 0.99$). Appendix A.5 provides information on the training and evaluation of raters.

3.2 Results

3.2.1 Experiment 1

Figures 3.2 and 3.3 show the results of Experiment 1, where participants in Generation 1 were novices. Over 10 generations, those who could observe the five models substantially improved in their image editing skills, in recreating the target image. Those who saw only one model demonstrated no significant improvement; if anything, they showed a decline in skill level. As the final row of Figure 3.3 shows, the *least* skilled learner in the 10th generation of the 5-Model treatment is superior to the *most* skilled learner in the 10th generation of the 1-Model treatment.

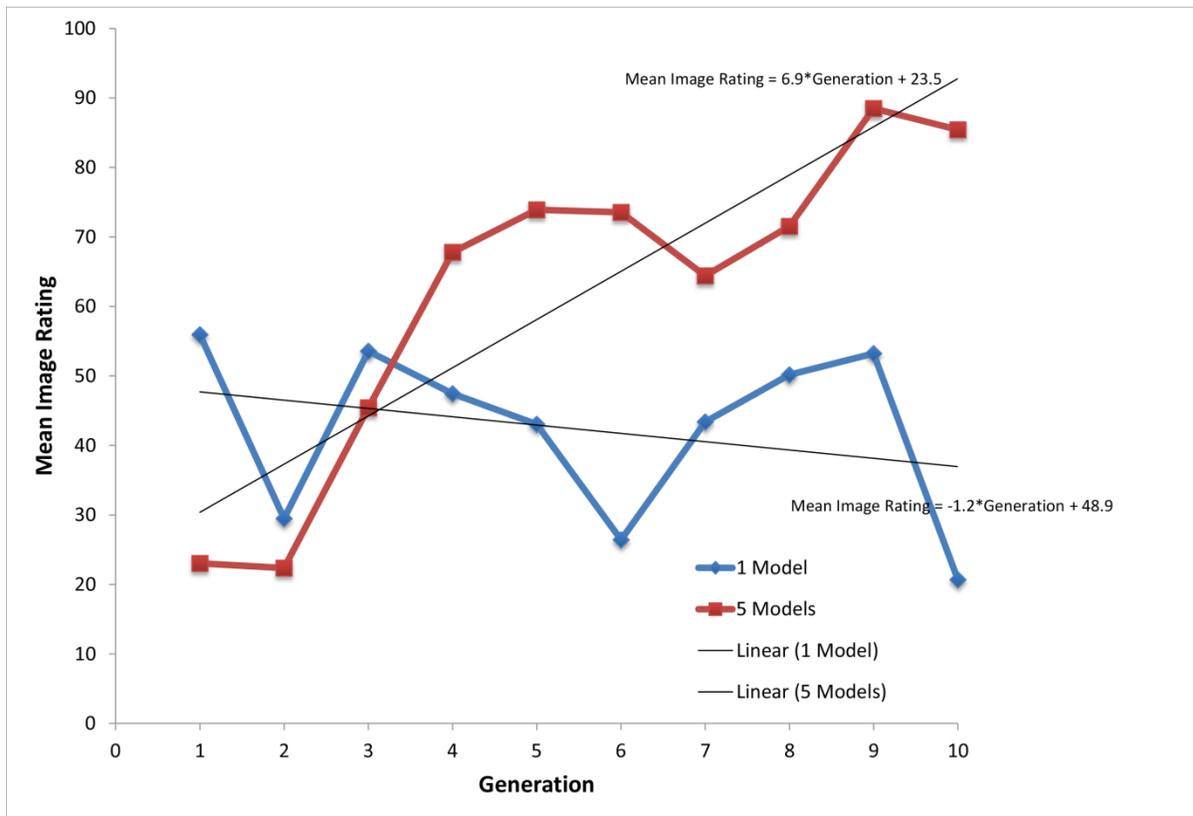


Figure 3.2 Mean Image editing skills over 10 generations for the 1-Model and 5-Model treatments in Experiment 1. Scores rescaled between 0 and 100, where 100 is a perfect score. Linear lines of best fit emphasize a cumulative improvement in the 5-Model treatment and no improvement, and a possible decline, in the 1-Model treatment.

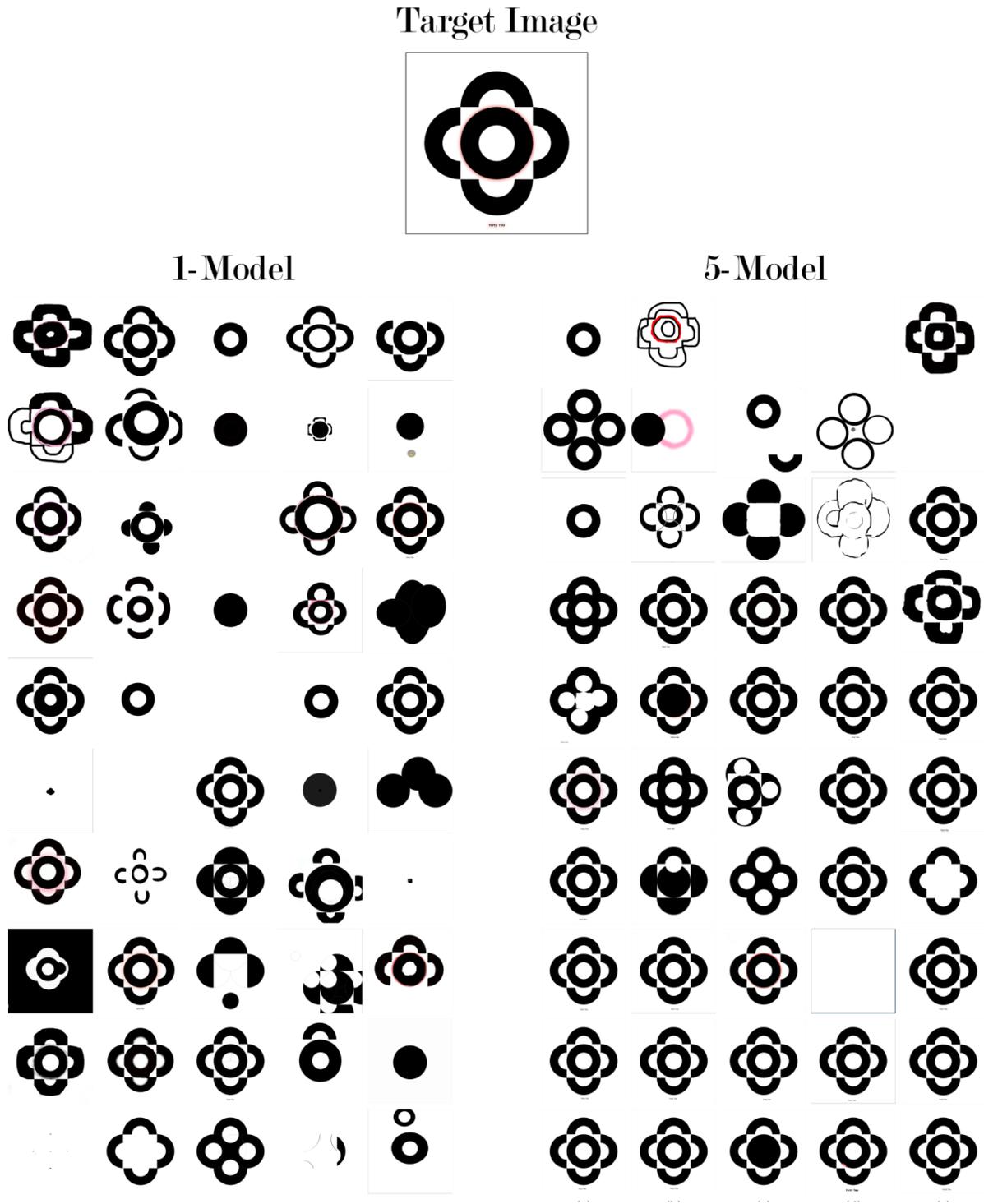


Figure 3.3 Experiment 1 final images from participants in the 1-Model and 5-Model treatments. The target image is included at the top for comparison. The columns are chains of participants in the 1-

Model treatment. Rows are generations going from top (Generation 1) to bottom (Generation 10). An obvious difference between the two treatments can be seen in the last row.

To further investigate the treatment differences visible in Figures 3.2 and 3.3, we regressed the standardized image rating scores on the main effects and interaction of generation number and treatment, controlling for Age and Male (gender with male = 1). Appendix Table A.3 contains the full series of regression models we examined. Of these models, the model controlling for Age and Male had the highest adjusted R^2 and is reported in Table 2.1, but the results are robust across all models. By alternating the dummy coding on treatment we are able to directly compare the effect of Generation on image rating score for each treatment. Our regression model (Table 2.1) estimated an average improvement of 0.23 standard deviations (equivalent to 7 percentage points) per generation in similarity-to-target image ($p < 0.001$), indicating the accumulation of skill. In contrast, there was only a small and non-significant effect of Generation in the 1-Model treatment, a decline of 0.06 standard deviations (2 percentage points) per generation ($p = 0.19$).

Table 3.1 OLS regression of standardized image rating scores on the main effects and interaction of Generation and Treatment (1-Model/5-Model), controlling for Male (gender, male = 1) and Age (standardized). By alternating the dummy coding of treatment, we directly compare the effect of Generation by looking at the Generation coefficients. In the 5-Model treatment, image ratings improve by 0.23 standard deviations per generation. In contrast, in the 1 model treatment there is no significant improvement in image ratings (and a possible decline).

	1-Model	5-Models
Generation	-0.06 (0.04)	0.23 (0.04)**
Age	-0.09 (0.08)	-0.09 (0.08)
Male	-0.36 (0.19)	-0.36 (0.19)
1-Model	-0.96 (0.36)**	
1-Model * Generation	0.28 (0.06)**	
5-Model		0.96 (0.36)**
5-Model * Generation		-0.28 (0.06)**
(Intercept)	0.11 (0.27)	-0.85 (0.26)**
R^2	0.353	0.353
Adj. R^2	0.319	0.319
N	100	100

** $: p < .01$, * $: p < .05$

Participants in the 5-Model treatment of Experiment 1 were given access to the images and notes from all five participants in the previous generation and could have learned from any or all of them. To examine selective learning biases, we broke down each participant's performance into 18 binary (present, absent) components, which gave us 810 (non-independent) observations for participants in generation 2 to 10. Then, using binary logistic regression, we regressed the presence or absence of each component in the participant's image on the presence or absence of each component in the participant's potential models, controlling for Age, Male and Generation. Each potential model was ranked from best (Model1) to worst (Model5). This allowed us to examine how participants weighted the relative importance of their potential models. We used clustered robust

standard errors (810 observations in 45 clusters) to control for common variance within each participant’s scores. The results (Table 3.2) indicate that the features present in the best model were the best predictor of the participant’s score. However, the 3 next best models were also predictive of participants’ scores, indicating that participants were also looking at other models. This suggests that participants were using a skill or success bias, with the greatest weight on the most skilled model, but with some non-zero weight on everyone else except for the worst model. Such patterns offer some evidence that participants were combining information from multiple models, thereby generating novel recombinations of elements not possessed by any single one of their teachers. Of course, given the error in the estimates, we can’t be too confident in the differences observed between models 2, 3 and 4.

Table 3.2 Binary logistic regression of the presence or absence of each component of the target image in each participant’s attempted image on the corresponding component in each of the 5 potential models. We control for non-independence between participant’s image components using clustered robust standard errors. The odds ratios reported reveal a large and significant bias for the best model, but also biases for the 3 next best models. We control for Generation, Male and Age.

Experiment 1	
Model1	3.910*** (1.258)
Model2	2.481*** (0.867)
Model3	1.747* (0.557)
Model4	2.187*** (0.583)
Model5	0.893 (0.260)
Pseudo R^2	0.283
N	810 (45 clusters)
Robust standard errors in parentheses	
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$	

3.2.2 Experiment 2

Figure 3.4 shows the results of Experiment 2, where participants in the first generation were knot-tying experts. The knot-tying skills of those in the 5-Model treatment decline more slowly than in the 1-Model treatment over the first three generations, and then level off to a higher average knot skill than those in the 1-Model treatment. Meanwhile, knot-tying skills in the 1-Model treatment continue to decline, though at a decelerating rate, through to generation 10.

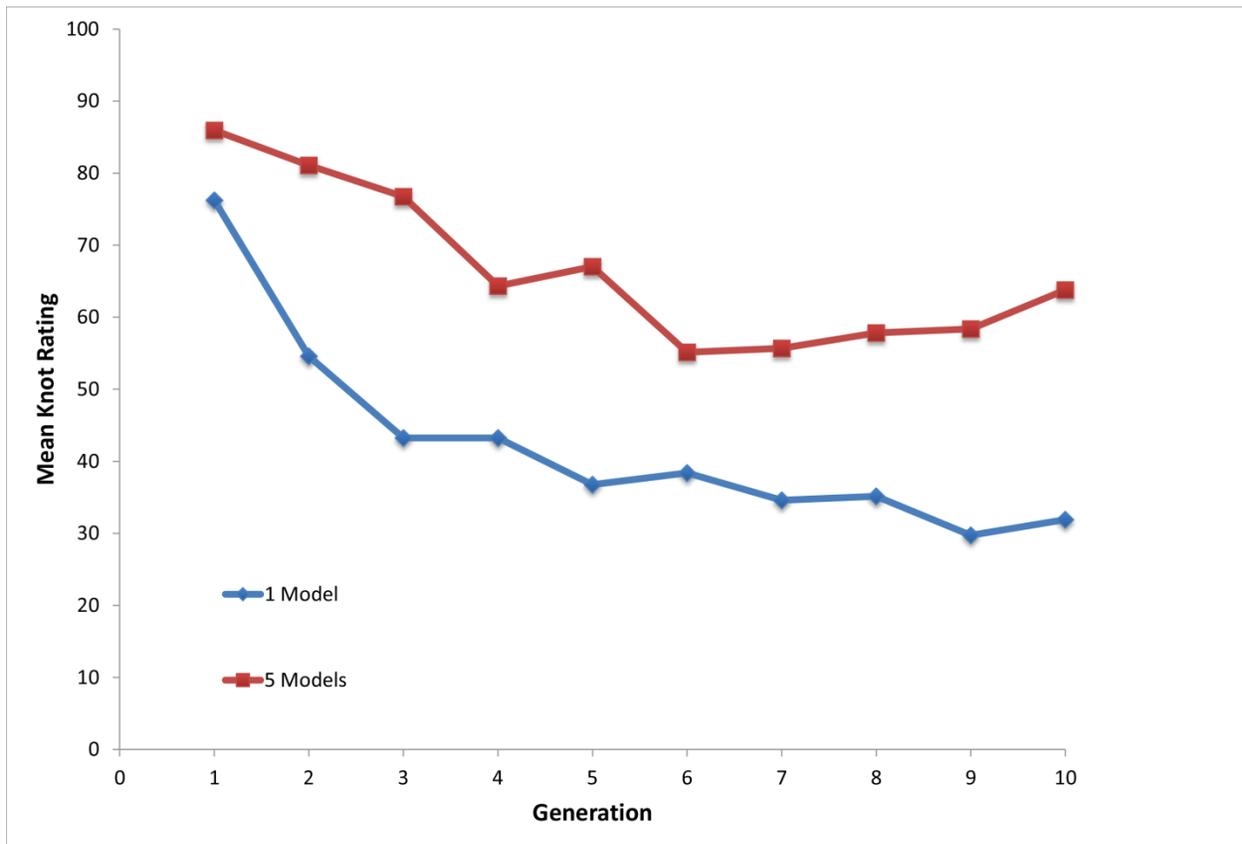


Figure 3.4 Mean knot-tying skills over 10 generations for the 1-Model and 5-Model treatments in Experiment 2. Scores rescaled to between 0 and 100, where 100 is a perfect score. The loss of skills is fastest in the first 3 generations and much faster in the 1-Model treatment than in the 5-Model treatment. Generations 4 – 10 suggest different equilibria where the 5-Model treatment has an equilibrium at twice the skill level of the 1-Model equilibrium.

To further investigate the difference between the treatments and generations in Figure 3.4, we separately estimated a series of OLS regression models for the first 3 and last 7 generations, controlling for Age, Male, Ethnicity, and experience with knot tying. Appendix Table A.5 contains the full series of regression models we examined. Of these, the models controlling for Age, Male, and knot tying experience had the highest overall adjusted R^2 values and are reported in Table 3.3, but results were robust across all models. Our regression model (Table 3.3) estimated that over the first three generations, the mean skill of the 5-Model treatment declines by 0.25 standard deviations (equivalent to 6 percentage points) per generation ($p = 0.37$) while the 1-Model treatment declines by 0.67 standard deviations (16 percentage points) per generation ($p = 0.02$).

From Generations 4 to 10, Table 3.3 shows that the mean skill in the 5-Model treatment declines at a rate of 0.03 standard deviations (0.6 percentile points) per generation ($p = 0.51$), while that in the 1-Model treatment declines at a rate of 0.07 standard deviations (1.2 percentile points) per generation ($p = 0.20$). While neither of these rates of loss is significantly different from zero at conventional levels, suggesting they may be approaching equilibrium, it is worth noting that the estimated magnitude of the rate of loss in the 1-Model treatment remains twice as large as that in the 5-Model treatment. And, a test of joint significance for the addition of the Treatment and Treatment-Generation interaction terms to model with only main effects reveals a significant increase in R^2 from 0.19 to 0.52, $F(62,64) = 21.6$, $p < .001$ (Appendix Table A.6).

Assuming Generation 10 is in the vicinity of the final equilibrium in skill, the mean skill level in the 5-Model treatment is twice that of the 1-Model treatment. In fact, every learner in the 10th generation of the 5-Model treatment is superior to the most skilled learner in the 10th generation of the 1-Model treatment.

Table 3.3 OLS regression of standardized knot rating scores on the Generation and Treatment (1-Model/5-Model), and their interaction, controlling for Male, Age (standardized) and knot-tying experience. By alternating the dummy coding of treatment, Table 3.3 directly compare the effect of Generation by looking at the Generation coefficients. The loss of skill within both the first 3 generations and the last 7 generations is twice as fast in the 1-Model treatment compared to the 5-Model treatment. We conducted a test of joint significance of treatment and treatment-generation interaction by statistically comparing regression models with and without these variables. Results indicate a statistically significant effect of treatment and treatment-generation interaction.

	1 Model:1-3	5 Models:1-3	1 Model:4-10	5 Models:4-10
Generation	-0.67 (0.28)*	-0.25 (0.28)	-0.07 (0.05)	-0.03 (0.05)
5-Model	0.28 (0.85)		0.78 (0.52)	
5-Model * Generation	0.42 (0.38)		0.03 (0.07)	
1-Model		-0.28 (0.85)		-0.78 (0.52)
1-Model * Generation		-0.42 (0.38)		-0.03 (0.07)
Age	0.07 (0.14)	0.07 (0.14)	-0.02 (0.09)	-0.02 (0.09)
Male	0.41 (0.37)	0.41 (0.37)	0.34 (0.18)	0.34 (0.18)
Sailing knot experience	0.11 (0.21)	0.11 (0.21)	0.01 (0.08)	0.01 (0.08)
Climbing knot experience	0.19 (0.33)	0.19 (0.33)	0.07 (0.07)	0.07 (0.07)
(Intercept)	1.40 (0.63)*	1.68 (0.65)**	-0.44 (0.38)	0.34 (0.38)
R^2	0.518	0.518	0.524	0.524
Adj. R^2	0.365	0.365	0.471	0.471
N	30	30	70	70

** $: p < .01$, * $: p < .05$

In Experiment 2, by contrast to Experiment 1, there was a substantial time cost to observing models, since participants could not watch all model videos in the available learning time and had to select fewer models from which to learn. Casual observations suggest that most participants watched only 1 video, or sometimes 2. For this reason, it's not clear what the relationship should be between the various models and the specific traits acquired by the learner, so we do not present analyses of this. Note that, although only 1 or 2 models were typically observed, these were typically the best models, replicating the dynamics of the simplest models (e.g. Henrich, 2004b).

3.3 Discussion

In a micro-society laboratory setting, our results confirm predictions made by existing formal cultural evolutionary models (Enquist et al., 2010; Henrich, 2004b, 2009a; Kobayashi & Aoki, 2012; Powell et al., 2009). Specifically, they confirm how increasing the number of accessible cultural models can generate greater accumulations of technical know-how in a population, such that every individual in the final generation of the 5-Model population is more skilled than the most skilled individual in the final generation of the 1-Model population and almost all individuals in the 1-Model population. The results confirm that more sociable populations can sustain more complex skills while less sociable populations gradually lose these skills over generations.

Our more detailed analyses of Experiment 1 indicate that learners in the 5 model condition learned, to at least some detectable degree, from the top four performers, though they did rely most heavily on the top performer among their cultural parents. This is important because, by drawing ideas, techniques and insights from different models, learners can end up with novel recombinations that none of their cultural parents possesses. This, in a sense, creates innovations without ‘invention’, ‘creativity’, or trial and error learning (Boyd, Richerson, & Henrich, 2011; Henrich, 2009b).

We chose to compare the 1-model and 5-model treatments for pragmatic reasons: 1 model represents the natural lower bound while 5 models provide a substantial increase in model number, giving us a the best chance to observed the predicted effect in a relatively small number of generations without escalating the learners’ costs of observing and evaluating a large number of models. We expect the effect of the number of models on skill level and evolutionary rate to show diminishing returns, limited by how much time participants have to evaluate and integrate culture

from multiple models and potential contribution of additional models. We could have just as easily used 2-4 models, though we expect that the effect sizes would have been a bit smaller.

Our findings also suggest why one prior study has failed to reveal any effects for model number on mean skill levels (Caldwell & Millen, 2010). The theoretical models we are testing predict that if some skill or other cultural trait is sufficiently easy to learn or cognitively transparent, then increasing the number of models available to learners will have little impact over generations on mean skill level or performance (Henrich, 2004b). Based on these theories and Derex et al.'s (2014) results, which tested the effect of task complexity, we suspect that the relative ease of the task used in Caldwell and Millen (2010)—making a simple paper airplane—was too easy to learn to observe the effects we found. Nevertheless, such findings address any concerns that our results were the inevitable consequence of the laboratory setup. Future research should examine a wider range of tasks, forms of transmission, and range of modeling treatments.

One concern with our setup is that our participants, motivated by money, were primarily concerned with acquiring the specific skills and techniques necessary to match an ideal type, embodied in our target image or the system of rock climbing knots. These tasks did not have any other immediate practical ends in themselves, such as hoisting a heavy object or communicating a message. While we think that varying the degree to which participants can focus on an immediate practical goal is well-worth exploring (Herrmann, Legare, Harris, & Whitehouse, in press), it's important to realize that many real and practical aspects of culture have the match-to-target format. For example, an Inuit making his first kayak has no chance of figuring out all the relevant engineering principles that are implicitly embodied in a good kayak, or of knowing the kayak's performance under the extreme conditions that he will encounter weeks or months later. But, he is likely to have another sturdy and well-performing kayak on-hand, to copy. Similarly, a !Kung hunter-

gatherer making his arrow poison using *Diamphidia* beetle larva, acaia sap, salvia and firing can only test his poison in real-time, while pursuing prey. Even then, the quality of his feedback on his poison's effectiveness will usually be murky. The best he can do in the short-term is follow the available recipe as closely as possible. We suspect functional end goals are mostly relevant for relatively easy tasks where individual learning can make a big difference.

Human and non-human primate populations vary in sociality. Chimpanzees and gorillas have mean group sizes of 51 and 7 respectively (Dunbar, 1992; Lind & Lindenfors, 2010), and interact only with their immediate group members. In contrast, although hunter-gather groups such as the Hadza live in camps of approximately 30 individuals (11.7 adults), such bands are embedded in much larger tribal networks (~500 adults; over 1000 individuals) comprising many camp sites, with whom they interact with extensively (Apicella, Marlowe, Fowler, & Christakis, 2012). Other hunter-gatherers have similar band sizes (e.g. !Kung, 23; Tiwi, 32; Mbuti, 104) and tribal networks (!Kung, 726; Tiwi, 2662; Mbuti, 1496) (Binford, 2001). Horticulturalists, such as the Yanomami, live in still larger villages of well over 100 individuals (Chagnon, 1988) with a total population of around 15,000. Understanding the relationship between sociality and cumulative cultural evolution is crucial to understanding the origins and ecological success of our species (Boyd et al., 2011; Dean et al., 2012; van Schaik et al., 2012). Several researchers have argued that cumulative cultural evolution, by giving rise to the skills and know-how related to complex tools, clothing, watercraft, fire, cooking, weapons, social norms and water containers, effectively drove our species genetic evolution over hundreds of thousands, if not millions, of years (Boyd & Richerson, 1988; Feldman & Laland, 1996; Gintis, 2011; Laland, Odling-Smee, & Myles, 2010; van Schaik & Burkart, 2011). If true, it is essential to explore how and why our lineage crossed the threshold into a regime of cumulative cultural evolution, but others did not. This study suggests that our sociality – our social networks,

conspecific tolerance, inter-group relations or population structure, may be what distinguished our ancestors from other primates, and pointed us on a different evolutionary trajectory (Hill et al., 2011).

Auspicious social conditions for crossing the cumulative cultural evolutionary threshold might emerge if ecological conditions caused a group living species, such as chimpanzees, to begin pair-bonding (Chapais, 2009). This could stimulate the emergence of (somewhat) peacefully interacting groups, which could increase the size and interconnectedness of populations, opening the door to the emergence of cumulative cultural evolution. Once the cumulative cultural evolutionary threshold is crossed, autocatalytic feedback between cultural learning, tool use and sociality may kick in to synergistically drive all three (Boyd et al., 2011; Chudek & Henrich, 2011).

Chapter 4: When and Who of Social Learning

Humans are a cultural species, heavily reliant on a rich repertoire of ideas, beliefs, values, and practices acquired from other members of their social groups. Evolutionary approaches to culture postulate that our species' social learning abilities – the psychological foundations that undergird these cultural repertoires – are genetically evolved cognitive adaptations for surviving in environments in which individually acquiring information is costly. Building on this, a large body of theoretical research has explored the conditions under which natural selection will favor various learning strategies (Boyd & Richerson, 1985, 1988, 1996; Henrich & Boyd, 1998, 2002; King & Cowlshaw, 2007; Nakahashi, Wakano, & Henrich, 2012; Perreault, Moya, & Boyd, 2012). This theoretical research provides clear predictions about when individuals, both human and non-human, should rely on their individual or asocial experience and when they should deploy one or more social learning strategies, such as conformist transmission (a tendency to disproportionately copy the majority or plurality). By contrast, relatively little empirical research has sought to directly test these models in the laboratory with human participants, though key exceptions with adult participants include McElreath, et al. (2005), Efferson, et al. (2008), and Morgan, et al. (2012) and with children include Wood, Kendal, and Flynn (2013), Haun, Rekers, and Tomasello (2012), Chudek, Brosseau-Liard, Birch, and Henrich (2013), and Morgan, Laland, and Harris (2014). Here, we aim to advance this research program empirically by testing some novel predictions and implications derived from existing theoretical work, as well as to replicate some prior results in new and more diverse populations. We test predictions regarding how (a) the number of cultural traits, (b) payoffs associated with different decisions, (c) fidelity of social transmission, and (d) group size influence the use of social over asocial learning, and the application of conformist biases within social learning. In addition, we consider the implications of existing models for predicting who might tend to use

which strategies, and use individual differences in cognitive abilities, social status, and cultural background to account for individual level variation in learning strategies (for a similar effort in other transmission contexts see Flynn and Whiten (2012)). Our efforts extend prior research on conformist biased social learning, which revealed much individual variation, but did not attempt to account for it.

4.1 Theoretical Research

Several evolutionary models (Boyd & Richerson, 1985, 1988, 1996; Henrich & Boyd, 1998) predict that reliance on social learning (over asocial learning) should increase with the cost or difficulty of asocial learning, the size of the majority, and the stability of the environment. These predictions make intuitive sense – individuals will prefer cheap, reliable, and accurate information; the reliability of social information increases with larger majorities and accuracy decreases with changes to the environment to which it pertains. Other models (King & Cowlshaw, 2007) predict that reliance on social learning should increase with access to more demonstrators, which typically increases with group size: More demonstrators reduce sampling error.

Within the realm of social learning, evolutionary models reveal the social learning strategies (Laland, 2004; Rendell et al., 2011) and biases (Boyd & Richerson, 1985) favored by different situations or circumstances. One such bias is conformist transmission. In a particular population, there may be many variants in behaviors, beliefs, or values, from herein referred to as *traits*. Conformist transmission (Boyd & Richerson, 1985) represents a type of frequency dependent social learning strategy in which individuals are *disproportionately* inclined to copy the most common trait in their sample of the population (e.g. individuals have a 90% probability of copying a trait that 60% of people possess). Conformist transmission is particularly important, because it tends to homogenize behavior within groups, increasing between group variation relative to within group variation (Boyd

& Richerson, 1985; Henrich & Boyd, 1998), strengthening the effect of intergroup competition on cultural variation (Chudek, Muthukrishna, & Henrich, 2015; Henrich, 2012), and potentially hindering cumulative cultural evolution within a group (Eriksson, Enquist, & Ghirlanda, 2007). Conformist transmission contrasts with unbiased transmission, whereby individuals copy a trait at the frequency found in the population (e.g. individuals have a 60% probability of copying a trait that 60% of people possess).

Several evolutionary models reveal the conditions when the conformist transmission bias is more adaptive than unbiased transmission. Typically, these models have analyzed only 2 traits. However, Nakahashi, Wakano, and Henrich (2012) have extended these models to N traits. Their model predicts that the strength of the conformist bias will increase with the number of traits in the environment. To understand the logic, consider a world with only 2 traits—black and white shirts. The presence of black shirts at anything above 50% suggests that people are selecting black shirts above chance. However, in a world with four traits – black, white, green, and red shirts – black shirts need only be present above 25% to suggest selection above chance. Thus, if 51% of people were clothed in black shirts, you would be much more likely to also wear a black shirt if there were 4 shirt options than 2 and even more so if there were 10 options and so on. One important implication of this model is that all current models and experiments may have been underestimating the strength of the conformist bias, because there are often more than 2 traits in the real world. In addition to the number of traits, the model also predicts that the strength of the conformist bias will increase with errors in transmission and with strength of selection⁹, consistent with other 2 trait conformist bias models (Henrich & Boyd, 2002). Other models (Perreault et al., 2012) predict that a stronger

⁹ We infer this last prediction based on migration less than 50% and weak selection (see Supplementary Materials of the published paper).

conformist bias will be more adaptive in larger groups, as information reliability increases, with an asymptotic relationship between group size and the strength of the conformist bias.

4.2 Experimental Research

In contrast to the growing body of theory, there has been relatively little experimental research investigating conformist biases. The first experimental test of these theories explored the effects of task difficulty and environmental variability (McElreath et al., 2005). The results revealed both unbiased and conformist transmission, with increased conformist transmission as the environment fluctuated. However, the results were inconsistent between experiments and were ultimately difficult to interpret. A later experiment by Efferson et al. (2008) separated participants into asocial and social learners and looked for evidence of a conformist bias among the social learners. On average, participants exhibited a conformist bias, but there was also considerable variation within participants, including some non-conformists. Most recently, Morgan et al. (2012) systematically tested nine theoretically derived hypotheses, including hypotheses related to group size, majority size, confidence, asocial learning cost and difficulty, number of iterations, participant performance, and demonstrator performance. In all cases, the results supported evolutionary predictions and found evidence of a conformist bias. All three sets of experiments described above revealed heavy reliance on social learning and the presence of a conformist bias, but they also documented, but did not explain, substantial individual variation. This individual variation has also been shown in social learning more generally (e.g. Whiten & Flynn, 2010). In the present research, we test several evolutionary theories and address this gap.

4.3 Present Research

In two experiments, we measure both the reliance on social learning and the strength of the conformist bias, testing several untested theoretical predictions. Based on the models, we predict

that reliance on social over asocial learning will increase with: (a) transmission fidelity (Boyd & Richerson, 1985, 1988, 1996; Henrich & Boyd, 1998) and (b) group size (King & Cowlishaw, 2007; Perreault et al., 2012). We predict that the strength of the conformist bias will increase with (a) number of traits (Nakahashi et al., 2012), (b) payoffs of the traits being copied (effectively the strength of selection; Nakahashi et al., 2012), and (c) errors in transmission (Henrich & Boyd, 2002; Nakahashi et al., 2012). Note that as transmission fidelity increases (i.e. errors in transmission decrease), reliance on social learning is expected to increase, but the strength of the conformist bias is expected to decrease. The decrease in the strength of the conformist bias with increased transmission fidelity may be more intuitive if you consider that the conformist bias helps to correct for errors in transmission. As errors increase, it pays to put more weight on larger majorities since they're less likely to emerge by chance. In testing these predictions, we also tested the effect of majority or plurality size in a more ethnically diverse population than past conformist transmission experiments.

We also developed and tested hypotheses to account for individual differences in social learning and conformist transmission. No work has yet shown what accounts for these differences, nor applied theoretical insights to understand the variation. Applying existing theory to individual variation, we explored three individual difference measures:

- a) **Cognitive abilities:** Individuals with better cognitive abilities ought to possess better private information, resulting in less individual uncertainty, which should result in reduced reliance on social learning and conformist transmission. Alternatively, those with better cognitive abilities may select the more adaptive strategy (i.e. copying when uncertain) – that is, cognitive abilities may in part be about selecting the best learning strategy overall.

- b) **Status:** Individuals who perceive themselves as higher in prestige status may reduce their reliance on learning from others who they perceive as less prestigious. Dominance status will bear no relationship to learning strategies once we control for prestige status and cognitive abilities.
- c) **Cultural Background:** Populations may differ in their tendency toward social learning and conformist transmission (Bond & Smith, 1996; Cialdini, Wosinska, Barrett, Butner, & Gornik-Durose, 1999; Mesoudi, Chang, Murray, & Lu, 2015). Cultural psychologists have argued that East Asians in particular are more likely to conform than Westerners. This may result in population-level differences in social learning and conformist transmission.

Besides these theoretically motivated variables, we also examined individual differences in (1) reflective thinking styles (intuitive vs reflective), (2) rule following, (3) personality, and (4) a variety of demographic variables.

4.4 Methods

We ran both our experiments on the same participants, but randomized the order of measures and experiments between groups. We report our participant demographics, general design, and specific procedures for each experiment.

4.4.1 Participants

We recruited 101 participants from the University of British Columbia's Economics Participant Pool, which is open to the public, but primarily consists of undergraduate students. Of these 101 participants, 27 participants failed at least one of our two vigilance check questions, leaving us with 74 usable participants (39 Female; Mean Age = 21.73, SD = 5.55). Including all participants is arguably defensible for our contextual variable analyses, because participants were incentivized for performance. Their inclusion generally strengthens our overall findings. However,

since these participants were not incentivized for completing the individual-difference measures and failed vigilance checks within them, we conservatively exclude them from the main analysis, but report all analyses with their inclusion in Appendix A.

4.4.2 General Design

We ran two experiments on all participants. In Experiment 1, we examined the effects of the number of traits. In Experiment 2, we tested the effects of payoffs and transmission fidelity. In both studies, we also explored group size (from 5 to 11 participants) and the proportion of people who selected each trait. In our experiment, *traits* are the lines of different length that participants selected between; we will refer to them as *options* from herein. As noted, we also measured several individual-level factors, detailed in Background Measures. Participants were paid a show-up fee of \$10 and could win an additional \$20 based on performance in the two experiments. Figure 4.1 illustrates the general design of the experiment.

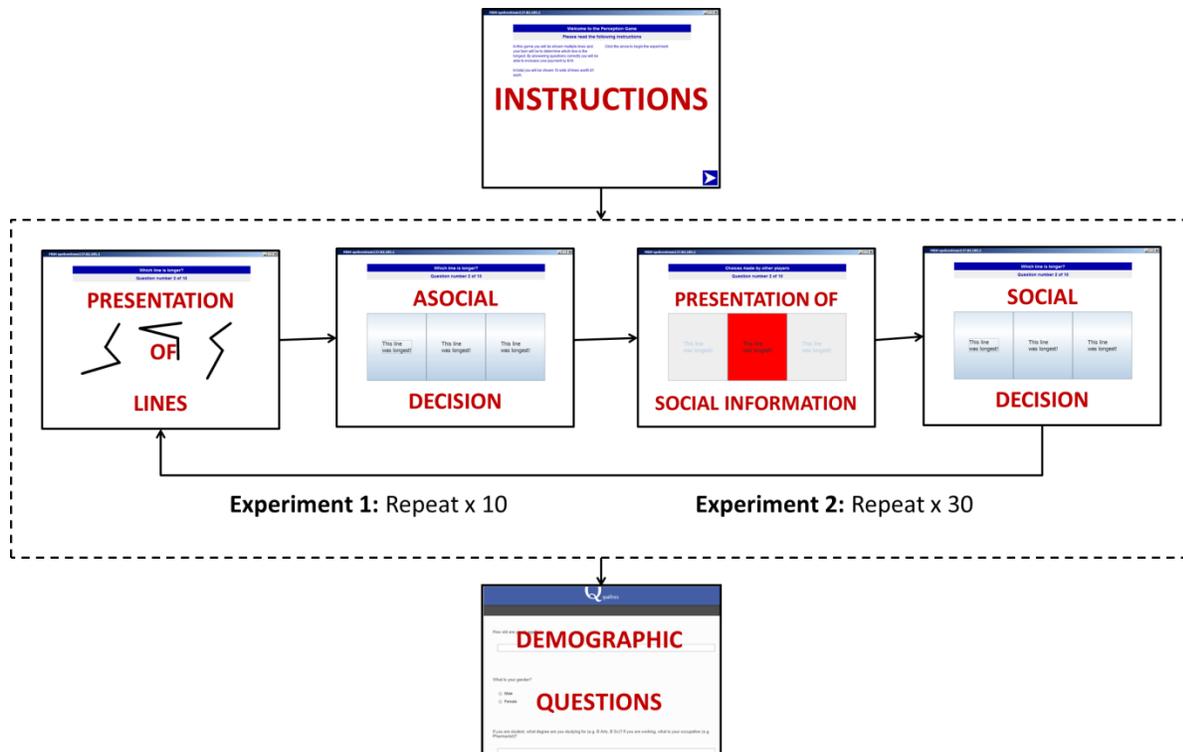


Figure 4.1 Flowchart of Experiment Design. The order of the experiments was randomized. We always asked demographic questions at the end, but we asked background measures (not shown) before or after all experiments (also randomized).

4.4.3 Experiment 1: Number of Options

In Experiment 1, participants had to compare between 2 and 6 lines to identify the longest line. This was repeated 10 times. The lines appeared for 3 seconds and then participants made their first ‘asocial’ decision. The software then displayed the decisions made by other participants one after another. The participants were shown panels corresponding to the different lines and each decision made by another participant was indicated by the corresponding panel flashing (red then gray). After receiving this social information, participants answered the question again. Keep in mind there was *no deception in this experiment*, so this was real social information.

Each trial was worth up to \$1. The payoff associated with each line was proportional to the length of the selected line relative to all other lines, with the longest line worth \$1, the shortest line

worth nothing, and lines of intermediate length worth a value less than \$1 based on the function graphed in Fig. S1 (see Appendix B.2 for details). With 10 trials each worth a maximum of \$1, participants could earn \$10 in this phase of the session. We informed participants at the beginning of the experiment that their payment depended only on their second response to each set of lines.

4.4.4 Experiment 2: Transmission Fidelity and Payoffs

In Experiment 2, we restricted the number of lines to 2 and varied the transmission fidelity and payoffs. The task involved comparing 30 pairs of lines to identify the longest line, with participants first giving an asocial response and then receiving social information and information about transmission fidelity before getting a chance to answer again. In other respects, participants went through the same process as in Experiment 1.

To explore the impact of transmission fidelity, we varied errors in transmission by replacing some of the social information with random computer generated answers. We informed participants of the probability of replacing real social information, which ranged from 0% (only true social information) to 40% (i.e. 60% social information, 40% random). See Appendix B.2 for a screenshot and details. After receiving this noisy social information, participants made their final decision.

To explore the impact of payoffs, we made the value of each trial between \$0 and \$2, with the ability to earn up to \$10 over 30 trials. The software clearly indicated the amount of money each question was worth before and throughout each trial.

We administered background measures either before or after the two experiments (randomly assigned with no significant difference between behavior or measures), but demographic questions (age, sex, time spent in Canada (some participants are immigrants), strategies used while playing the game, etc.) were always asked at the end.

4.4.5 Background Measures

Our three key individual-difference predictors were:

- **IQ:** We measured IQ using Raven's Advanced Progressive Matrices (Raven & Court, 1998).
- **Prestige and Dominance:** We measured self-reported prestige using the Prestige and Dominance scale (Cheng, Tracy, & Henrich, 2010).
- **Cultural Background:** We asked for participant ethnicity, if they had lived their entire lives in Canada, how well they speak their native language, how much they identify with Canada (Inclusion of Other in the Self Scale; Aron, Aron, & Smollan, 1992), and their degree of acculturation (Vancouver Index of Acculturation; Ryder, Alden, & Paulhus, 2000).

To pre-emptively counter other potential explanations for variation in social learning and conformist transmission, we also measured:

- **Reflective vs Intuitive Thinking Styles:** We measured reflective vs intuitive thinking styles using the Cognitive Reflection Test (CRT; Frederick, 2005). We included the CRT since it is plausible that copying or not copying others may be an intuitive decision. In this case, intuitive or reflective thinking styles will predict social learning and conformist transmission.
- **Rule Following:** We measured the tendency to follow rules using the Rule Following Task (RFT; Kimbrough & Vostroknutov, 2013). We included the RFT since it is plausible that copying or not copying simply represents the rule in our experimental setting, in which case the tendency to follow rules will predict social learning and conformist transmission.

Finally, we included age, sex, and the Big 5 Personality Inventory, which are often a source of individual-differences. Further details can be found in Appendix B.4.

4.5 Analysis

Our first theoretical question concerns how our contextual variables influenced social learning and conformist transmission. In our analysis of social learning, we looked at the proportion of times participants changed their decision after viewing social information for each level of our predictor variables. We graphed these relationships and described them with a best-fitting function, and then predicted this binary decision (changed vs did not change) using our predictor variables. This analysis allowed us to look at how our manipulated predictors affected the use of social information, but we could not use the proportion of participants as a predictor, since those in the majority or plurality would themselves be less likely to change their decision.

To address the question of how majority size affected social learning with 2 traits, we followed Morgan et al. (2012): Participants are considered to have used social information if (a) their decision after viewing social information differed from their asocial decision *and* (b) the majority of other participants disagreed with the participant's original decision. In Experiment 1, there were pluralities rather than majorities (multiple options), and there was more information (e.g. relative proportions), which participants may have incorporated in addition to just the overall plurality. Here, we analyzed the data with all responses (not just where the plurality disagreed with the participant), but focused on the cases where participants changed their decision. In each case where a decision was changed, we looked at the frequency of each option; the frequency of the options the participants ultimately selected and the frequency of the options the participants did not ultimately select.

Finally, to determine the strength of any conformist bias, we ran an analysis where we calculated a single best-fit conformist transmission parameter (α) by aggregating the data across all individuals for each level of our key predictors – number of options, transmission fidelity, and payoff value – except group size, where we did not have enough participants in each level. To accomplish this, we used a Signal Detection Theory (SDT) perspective, considering the four possible decision scenarios for a particular option and frequency. Note that this is for each particular option. To illustrate, we use Line 2 (of between 2 and 6 lines) as the particular option:

SDT 1. Choosing the option both asocially (before seeing social information) and socially (after seeing social information). E.g. Line 2 is selected before seeing social information and Line 2 is selected again after seeing social information.

SDT 2. Choosing the option asocially, but choosing a different option socially. E.g. Line 2 is selected before seeing social information, but a different line (not Line 2) is selected after seeing social information.

SDT 3. Choosing a different option asocially, but choosing the option socially. E.g. a different line (not Line 2) is selected before seeing social information, but Line 2 is selected after seeing social information.

SDT 4. Choosing a different option asocially and socially. E.g. a different line (not Line 2) is selected before seeing social information and a different line (not Line 2) is selected after seeing social information.

In SDT 1, we have no way of assessing if a decision was based on the social information or asocial prior. In contrast, in the other three cases, we know that the proportion was insufficient to retain the decision (SDT 2), the proportion was sufficient to make them choose the option (SDT 3), or the proportion was insufficient to make them choose the option (SDT 4).

We used a logistic function to fit a sigmoid to these latter three cases (SDT 2-4), similar to earlier theoretical work in social learning (McElreath et al., 2008; Szabó & Tóke, 1998; Traulsen, Pacheco, & Nowak, 2007):

$$p_i = \frac{1}{1 + e^{-\alpha(p_t - c)}}$$

Where p_i is the probability of choosing option t and p_t is the frequency of option t . The α parameter of the sigmoid is a measure of the strength of the conformist bias. If $\alpha < 0$, this indicates anti-conformity and if $\alpha \approx 0$, we assume decisions are being made independent of social decisions, i.e. no social learning. In contrast, $\alpha < 5$ suggest some social learning, but not conformist transmission. Finally, $\alpha \geq 5$ is evidence of conformist transmission, with higher values indicating a stronger conformist transmission bias. The c parameter tells us the inflection point, i.e. when individuals are 50% likely to choose the option and suggests a conformist bias when $c < 0.5$. These four categories match four types of formally defined frequency-dependent social learning strategies, which we discuss in Appendix B.1.

Nakahashi, et al. (2012) predict that c should be inversely related to the number of options (N), i.e. $c = 1/N$ – this is the frequency at which the trait would be present at chance levels. We used a nonlinear least-squares (NLS) estimate to fit α and c in Experiment 1 with multiple options, measuring the strength of the conformist bias and testing Nakahashi et al.’s (2012) theoretical predictions. In Experiment 2, with only 2 options, we set $c = 0.5$, the expected inflection point ($c = 1/2$) to fit the strength of the conformist bias (α). In Figure 4.2, we plot the sigmoid based on this function for different values of α and c .

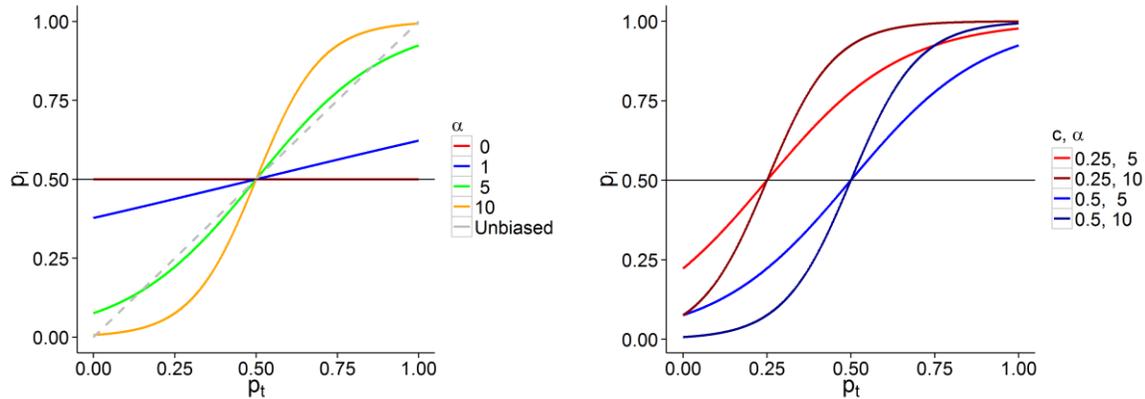


Figure 4.2 Logistic function sigmoid for different values of α (with $c = 0.5$ on left) and different values of c (right). The α parameter determines the curvature of the sigmoid and therefore the strength of the conformist transmission bias. The c parameter determines the inflection point.

Our second theoretical question was what individual factors predicted the strength of conformist transmission. To answer this second question, we fit the strength of conformist transmission to all responses for each individual separately. We then regressed these individual-level α values on our individual-level predictors.

4.6 Results

We report the results for contextual predictors and then individual predictors, analyzing Experiment 1 and 2 separately. We analyze the effect of each predictor on social learning and then the strength of the conformist bias.

4.6.1 Number of Options (Experiment 1)

Recall that in Experiment 1 participants had to select the longest line from between two and six options. We begin by analyzing the effect of the number of options on people's reliance on social learning over asocial learning.

4.6.1.1 Social Learning

Figure 4.3 shows a non-linear relationship between the number of options and the percentage of decisions that changed after seeing social information. With only 2 options, a little over 10% of people changed their decision after viewing social information, but this number rises to over 25% with 4 options and to almost 30% with 6.

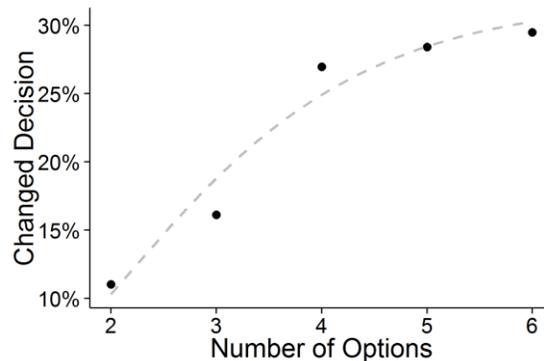


Figure 4.3 Percentage of decisions that were changed after seeing social information for different number of options. Although there are too few points to be certain about the function that best fits these data, we used a non-linear least squares method to fit to the reciprocal of the number of traits ($y = -0.60 \frac{1}{n} + 0.40$), plotted with a grey dashed line. Our choice of fitting the reciprocal of the number of traits is based on the logic underlying the Nakahashi, et al. (2012) model i.e. the probability of selecting the trait at chance is $\frac{1}{n}$.

Next, we look at how the frequency of each option in the social information predicted changing to that option. To do this, we use a binary logistic model to regress participant's decisions on the proportion of participants who selected an option (*Proportion*), the number of options (*Options*), and number of participants in the group (*Participants*), thereby testing several theoretical predictions (Boyd & Richerson, 1985, 1988, 1996; Henrich & Boyd, 1998; King & Cowlshaw, 2007). Each participant made multiple decisions. We control for common variance created by multiple observations from the same person with random effects for each individual. We remove

age and gender from the analysis; neither was significantly predictive and made very little difference to the results (see Appendix B.4 for full models). Nakahashi et al. (2012) made no specific predictions about the functional form of the relationship between the rate of social learning and number of traits. But, guided by their predictions for the conformist bias and predictions made by other models for the effect of the cost of asocial learning (which should increase with more traits), we test a model with the number of options (Model 1) and a model with the reciprocal of the number of options ($1/(N - 1)$; Model 2). We report these in Table 4.1.

Table 4.1 Binary logistic multilevel model of decision to switch regressed on the proportion of participants in the option (in 10% increments for easier interpretation), the reciprocal and number of options (separate models), and the number of participants in the group. All coefficients are odds ratios. We control for common variance created by multiple observations from the same person with random effects for each individual.

	Model 1	Model 2
	Number of Options	Reciprocal of Options
(Intercept)	<0.01*** [0.00,0.01]	0.05** [0.01,0.42]
Proportion 10%	3.62*** [2.82,4.87]	3.56*** [2.79,4.78]
Options	1.68** [1.23,2.35]	0.04** [0.01,0.32]
Participants	0.96 [0.76,1.21]	0.92 [0.73,1.16]
AIC	171.67	173.13
Obs.	332	332
Groups	64	64

***: $p < .001$, **: $p < .01$, *: $p < .05$

Table 4.1 reveals that participants are much more likely to change their decision overall if there are more options – 1.68 times as likely for every additional option. Participants are also more likely to change their decision as the proportion of others who select the option increases – 3.6 times as likely for every additional 10% of participants. Our results indicate that the number of

participants in the group (5-11) did not affect the likelihood of changing the decision. Based on the AIC values (Table 4.1) the fit of number of options and reciprocal of options models were almost identical.

4.6.1.2 Conformist Bias

To examine the influence of multiple options on the strength of the conformist bias in social learning, we fit the logistic function described in the Analysis section to the frequencies participants saw and their decisions for each number of options. We did this by combining all participants for each level of options – 2, 3, 4, 5, and 6. Thus, for each number of options, we calculate the strength of conformist bias (α) and the inflection point (c), i.e. what percentage of demonstrators need to have selected an option for the participant to copy that option with a 50% likelihood.

Figure 4.4a reveals that with each additional option, the strength of the conformist bias increases, but consistent with Nakahashi et al. (2012), the size of each increase decreases. Figure 4.4b reveals that the inflection point decreases reciprocally with increasing options, as predicted by Nakahashi, et al.'s (2012) model, though the actual value is higher than theoretical predictions (shown as a solid line to distinguish it from dashed lines fitted to the data). The difference between the experimental measurements and theoretical prediction may be an indication of the size of participants' asocial prior, which Nakahashi, et al.'s model does not address – they model a situation where individuals can only access either asocial or social information (but not both). The pattern in Figure 4.4b is what one would expect if individual's can combine asocial and social learning, as is the case in our experiments.

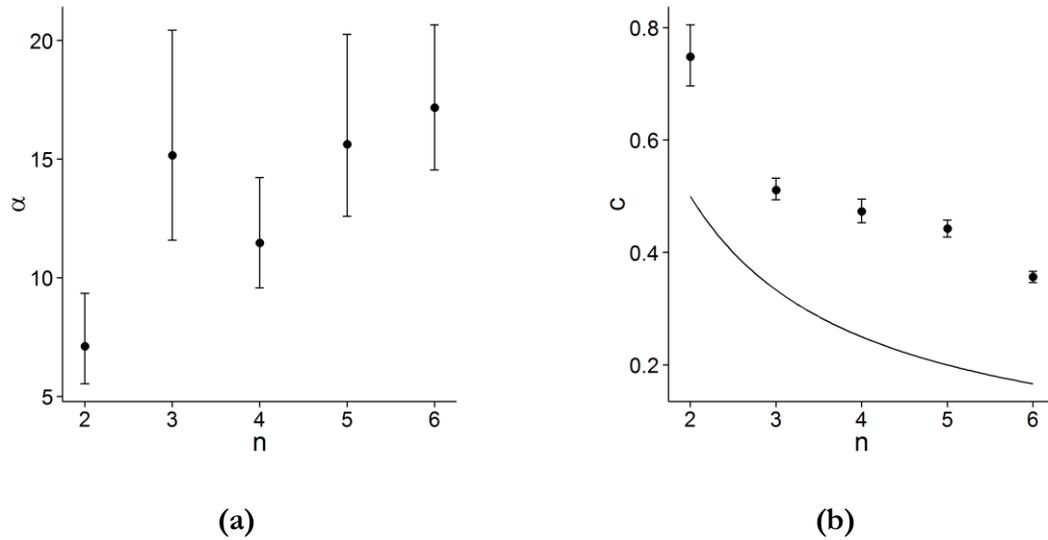


Figure 4.4 (a) Strength of conformist transmission parameter (α) as a function of number of options (n). The strength of the conformist transmission bias increases with more options. (b) Inflection point of logistic function as a function of number of options. The predicted value based on Nakahashi, et al (2012) is shown as a solid line. The inflection point decreases, but remains higher than the predicted value, indicating an asocial prior.

Figure 4.4b reveals the point at which individuals will select an option 50% of the time (c). With only 2 options, individuals select an option 50% of the time if 75% of others select it. With 4 options, individuals select an option 50% of the time if 50% of others select it. And with 6 options, individuals select an option 50% of the time if just 35% of others select it. Figure 4.4a reveals a measure of the gradient of the sigmoid (α). To get a sense for what these two parameters are telling us, consider what happens when someone sees 80% of other people select an option. If there are 2 options ($\alpha = 7$ and $c = .75$), the person has a 59% probability of changing their decision, but if there are 6 options ($\alpha = 17$ and $c = 0.35$), the person has a 99.95% probability of changing their decision. Together, these results reveal that as the number of traits in an environment increases, both social learning and the strength of the conformist bias increase, but at a diminishing rate.

4.6.2 Transmission Fidelity and Payoffs (Experiment 2)

Experiment 2 varied errors in the transmission channel and payoffs. To remain consistent with most existing theoretical models and with prior experimental research, we restricted choices to 2 options (instead of the 2 to 6 options in Experiment 1). As for Experiment 1, we first examine how these 2 factors influence social learning, and then look at their effect on the strength of the conformist bias.

4.6.2.1 Social Learning

Reliance on social information increased with higher fidelity transmission. Figure 4.5a suggests a linear relationship between transmission fidelity and the percentage of decisions that changed after seeing social information. At 100% transmission fidelity, about 16% of people changed their decision after viewing social information, but this number drops to 11% at 60% fidelity. Though this increase with fidelity is consistent with theoretical expectations, the differences in social learning were small; participants were not particularly responsive to our rather explicit manipulation of transmission fidelity.

Reliance on social information increased between having no payoff and some payoff, but did not increase with higher payoffs. Figure 4.5b shows that the percentage of decisions that changed after seeing social information increased by about 3% in moving from a zero payoff to 10 cents, but then remained consistent between 13% and 15% up to payoffs of \$2. The difference between zero and even a small payoff is consistent with prior experimental work on the Zero Price Effect (Shampanier, Mazar, & Ariely, 2007). One possible explanation for the lack of effect of increasing payoffs is that our experiment did not have the range or sensitivity to capture the effect of payoffs. For the transmission rates used in our experiment, Nakahashi et al. (2012) predict small and diminishing returns for low payoffs (weak selection in the model).

As in Experiment 1, we use a binary logistic multilevel model to regress participant decision on the size of the majority, transmission fidelity, question payoff, and number of participants in the group. We control for common variance created by multiple observations from the same person with random effects for each individual. We removed age and gender from the analysis; neither was significantly predictive and made very little difference to the results (see Appendix B.4 for full models). We consider majority percentage and transmission rate in 10% intervals and payoffs in 10-cent intervals for more intuitively interpretable coefficients (Model 1). We also ran a second model with payoffs as a binary variable with no payoffs vs non-zero payoffs (Model 2).

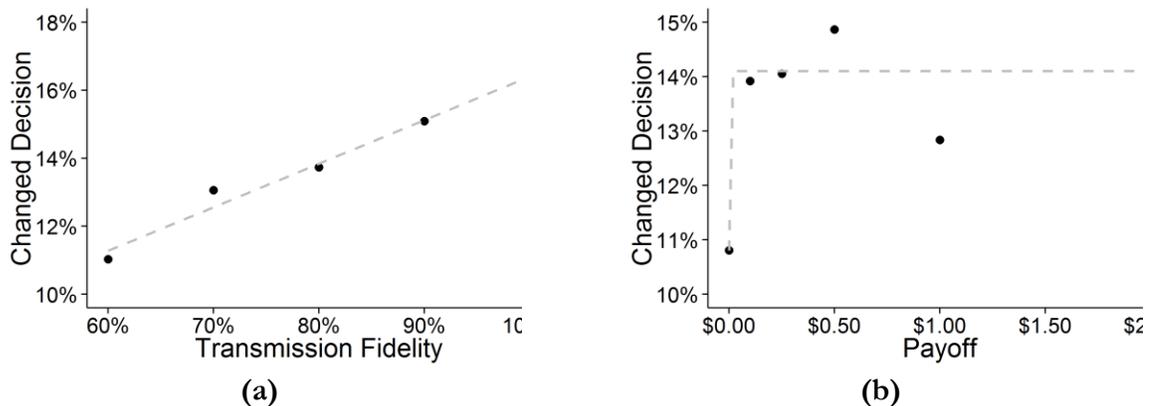


Figure 4.5. Percentage of decisions that were changed after seeing social information for (a) different levels of transmission fidelity, and (b) different question payoff values. Although there are too few points to be certain about the function that best fits these data, we used a non-linear least squares method to fit (a) to a linear model ($y = 0.13x + 0.04$), and (b) to a step-function ($y = 0.14$ if $x > 0$; $y = 0.11$ if $x = 0$). Fit functions are plotted with a grey dashed line.

Table 4.2 reveals a large effect of majority percentage, such that every 10% increase is associated with participants being 3.5 times more likely to change to the majority. We also find a large positive effect of transmission fidelity, with every additional 10% increase in fidelity associated with participants 1.3 times as likely to change to the majority. Consistent with Figure 4.5b, we see no linear effect of payoff, but a significant difference between zero payoff and non-zero payoffs

(participants are 2.6 times as likely to switch to the majority with some payoff). Finally, every additional participant in the group results in participants 1.28 times as likely to switch to the majority. Except for payoffs, these results are consistent with our theoretical predictions (Boyd & Richerson, 1985, 1988, 1996; Henrich & Boyd, 1998; King & Cowlshaw, 2007).

Table 4.2 Binary logistic multilevel model of decision to switch to majority on majority size, transmission fidelity, payoff, and number of participants in the group. All coefficients are odds ratios. We control for common variance created by multiple observations from the same person with random effects for each individual.

	Model 1 Linear Payoff 10c	Model 2 Binary Payoff
(Intercept)	1.22 [0.81,1.84]	0.49 [0.16,1.21]
Majority 10%	3.50*** [2.84,4.32]	3.60*** [3.06,5.01]
Transmission 10%	1.29*** [1.12,1.49]	1.31*** [1.13,1.55]
Payoff	1.01 [0.97,1.06]	2.62* [1.16,7.79]
Participants	1.28** [1.07,1.53]	1.28* [1.09,1.59]
AIC	761.47	750.10
Obs.	818	818
Groups	74	74

***: $p < .001$, **: $p < .01$, *: $p < .05$

4.6.2.2 Conformist Bias

To analyze the effect of the number of options on the strength of the conformist bias in Experiment 1, we fit the logistic function described in the Analysis section for 2 options, 3 options, and so on. Here, in Experiment 2, we perform the same analysis for each level of transmission fidelity (60%, 70%, 80%, etc.) and then each level of payoffs (0c, 10c, 25c, etc.).

Transmission fidelity significantly increases the strength of the conformist bias between 60% and 70% fidelity, but there is no clear difference above 70% (see Fig. 6a). Recall that in contrast, social learning increases linearly with transmission fidelity. The difference in the strength of the

conformist bias between 60% and 70% fidelity is large. An individual who sees 80% of others select an option will be 85% likely to copy that option if transmission fidelity is 60%, but will be 95% likely to copy the option if transmission fidelity is 70%.

Higher payoffs predict a stronger conformist bias (although the large confidence intervals make it difficult to determine if this trend is more than chance; see Figure 4.6b). The very large confidence interval on \$1 and \$2 may be due to fewer cases for these values. To compensate for this, we averaged the \$1 and \$2 cases in Figure 4.6c. These results suggest that higher payoffs lead to a stronger conformist transmission bias, with diminishing returns. Recall that we saw no trend in social learning, except between no payoff and some payoff. Thus payoffs have little effect on social learning, but do have an effect on the conformist social learning bias. Overall, these results only partially support the theoretical predictions. We will return to this in the Discussion.

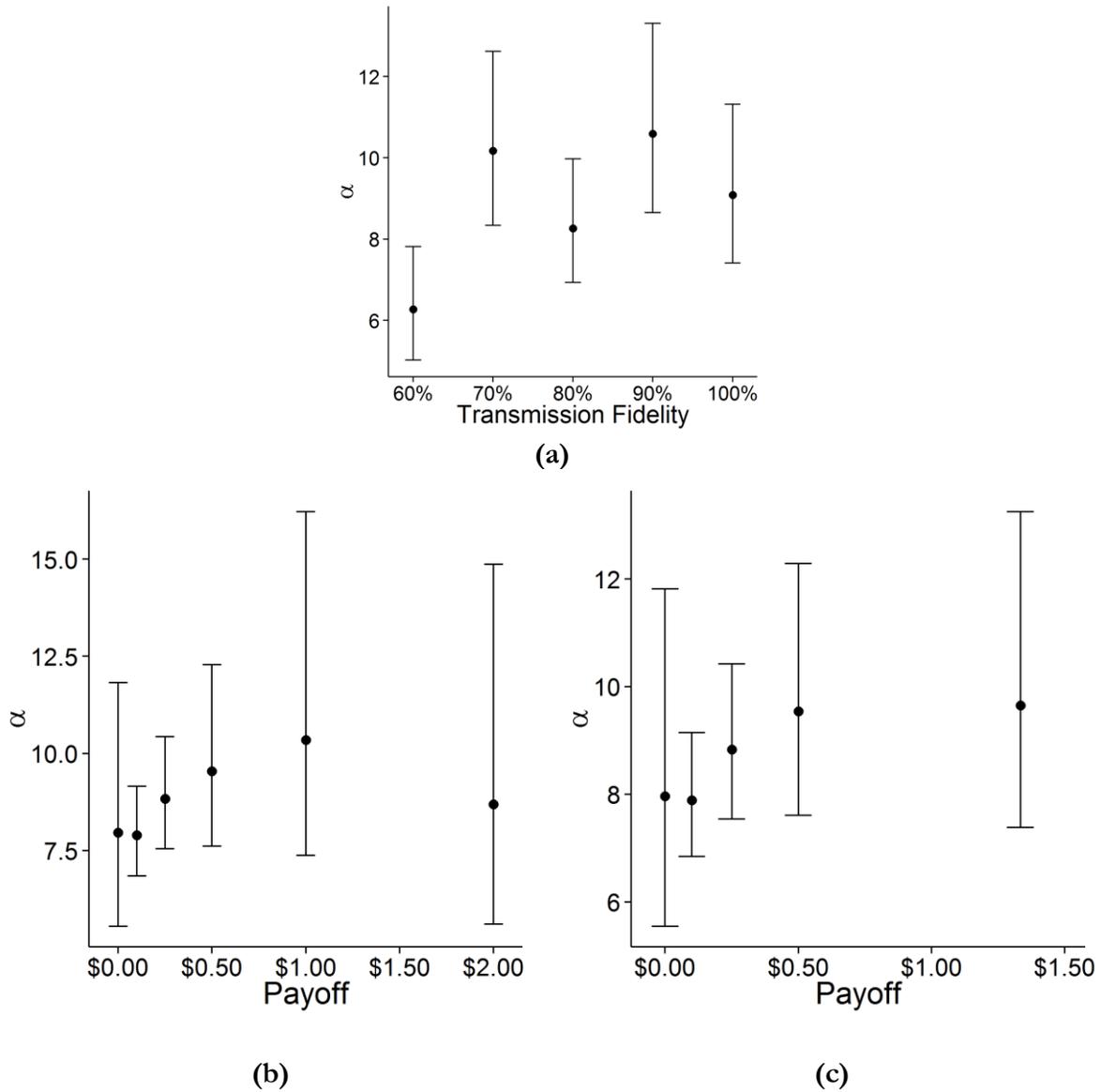


Figure 4.6 (a) Strength of conformist transmission parameter (α) as a function of transmission fidelity. Conformist transmission is strong when fidelity is higher than 60%, but at 60% it's only slightly above unbiased transmission. Strength of conformist transmission parameter (α) as a function of question payoff with (b) all payoff values and (c) \$1 and \$2 averaged to increase sample size for the highest value. The strength of the conformist transmission bias increases with diminishing returns as the payoffs increase.

4.6.3 Individual Variation in Social Learning Strategies

Consistent with past empirical research (Efferson et al., 2008; McElreath et al., 2005; Morgan et al., 2012), we found evidence of substantial individual variation in social learning and social learning strategies. We used the same analytic approach as in the previous sections analyzing social learning and then conformist transmission. To measure reliance on social information, we calculated the percentage of decisions that each participant changed after seeing social information. To measure the strength of the conformist bias (α_i), we fit a logistic curve based on the frequency of options they saw. We then regressed the social learning measure and the conformist bias measure on our theoretically motivated predictors (IQ, prestige, and culture), as well as several other measures that have been used in the literature, including reflective thinking styles, rule following, personality, and a variety of demographic variables.

4.6.3.1 Social Learning

In both experiments, IQ was significantly predictive of lower reliance on social information (see Table 4.3). Every standard deviation increase in IQ¹⁰ resulted in a 4% reduction in social learning in Experiment 1 and a 2% reduction in social learning in Experiment 2. This effect is small, but reliable.

¹⁰ Note: This standard deviation refers to a standard deviation within our experiment rather than IQ normalized to the population.

Table 4.3 OLS regression model percentage of decisions that were changed after viewing social information regressed on theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the reference group. These results show a negative relationship between IQ and social learning with higher IQ resulting in less social learning. The regression models reported show all theoretically inspired predictors; the regression model is significant when the non-significant predictors are removed (see Appendix B.4).

	Experiment 1	Experiment 2
(Intercept)	0.22 (0.15,0.30)***	0.20 (0.02,0.39)*
zIQ	-0.04 (-0.08,-0.00)*	-0.02 (-0.05,-0.00)*
zPrestige	-0.01 (-0.05,0.02)	-0.01 (-0.05,0.02)
East Asian	0.00 (-0.09,0.09)	-0.01 (-0.06,0.04)
Other Ethnicity	-0.02 (-0.12,0.08)	0.03 (-0.03,0.10)
zAge	-0.00 (-0.04,0.04)	-0.00 (-0.03,0.02)
Male	0.01 (-0.06,0.09)	-0.00 (-0.05,0.04)
R^2	0.09	0.09
N	74	74

***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$

With the exception of IQ, no other predictors were reliably predictive. Neither prestige nor cultural background were sizably or significantly predictive. Nor were other plausible predictors such as reflective thinking styles, rule following tendencies, personality, dominance, and a variety of other demographic variables. Since our participants included partially acculturated individuals, we attempted to predict social learning using the interaction of cultural background and measures of acculturation and cultural identification. These were also not sizably or significantly predictive. However, all of our predictors together account for only about 9% of the variance in social learning.

We briefly return to our null results with regard to cultural differences and prestige in the Discussion.

4.6.3.2 Conformist Bias

To assess the variation in the strength of conformist biases in social learning, we fit a logistic curve to all participant responses in Experiment 1 and 2 separately, assuming an inflection point of $1/N$, in order to fit the model. For all models, we again used a SDT approach, focusing on the 3 cases of interest and used the NLS method to estimate parameters.

Figure 4.7 Density distribution of α conformist transmission values in (a) Experiment 1 and (b) Experiment 2, with α calculated after scaling frequency of options by transmission fidelity. The red line indicates the cut off for conformist transmission with values to the left of this line indicating unbiased social learning. The x-axis is log-scaled. For visualization purposes, we removed some outliers – see Appendix B.5 for figure including these.

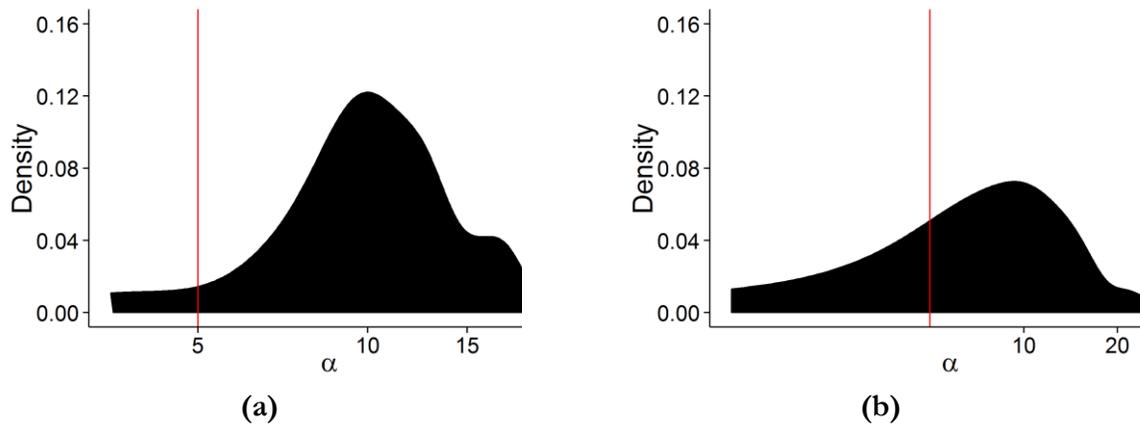


Figure 4.7 shows the distribution of α_i values in both experiments, with the vertical line marking unbiased, as opposed to conformist, transmission. In Experiment 1, only 3% of people showed unbiased social learning (or weaker). The remaining 97% of participants showed a conformist transmission bias to varying degrees, with the modal value a bit above 10. We found no

evidence of anti-conformity. In Experiment 2, 15% of participants showed unbiased social learning (or weaker) when data was fitted to the raw majority percentage. However, this value may inflate the tendency toward unbiased social learning because it combines individuals relying on social information with very different transmission fidelities. To address this, we scaled the majority size by the transmission fidelity and re-estimated α_i . With this adjustment, the percentage of unbiased social learners dropped to 9%. The remaining 91% of participants, or 85% for the unscaled calculation, showed some conformist transmission bias, with a modal strength close to 10. These results further support the argument that fewer options underestimate the strength of the conformist transmission bias. In neither experiment did we find any evidence of anti-conformity (Morgan et al., 2014)—negative α_i values.

In Table 4.4, we regress the strength of the conformist transmission bias on our theoretically inspired individual predictors. Because the distribution of the α parameter was highly positively skewed, we took the logarithm of this value before standardizing it (see Figure 4.7). For Experiment 2, we used the scaled α_i values, in part because it resulted in a better fitting model. However, no substantive differences were found using the unscaled fitted values, reported Appendix B.5.

Unlike our analysis of social learning above, the regression models in Table 4.4 reveal that the conformist bias is higher among those with low IQs *and* those with high IQs, compared to more average individuals. We found these results in both Experiments 1 and 2. We also found that the conformist bias was stronger in females and increased with age. Females had α_i values half a standard deviation higher than males, which translates to $\alpha_i = 1.6$ higher. For age, every 5.6 years translated to an $\alpha_i = 1.5$ increase. However, we had a limited age range with a mean age of 22. These differences were only found in Experiment 1, which is arguably more sensitive than Experiment 2, because there are often more than 2 options.

As with social learning, other analyses revealed no effect of the other plausible predictors and no effect of increased acculturation or identification. Note that, unlike with social learning, we had no specific predictions about the effect of social status (prestige or dominance) on conformist transmission.

Table 4.4 OLS regression model of standardized log measures of strength of conformist transmission (α) regressed on our theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the reference group. These results suggest a consistent quadratic (U shaped) relationship between IQ and the strength of the conformist transmission bias. Both those who scored high or very low on the IQ test were more likely to have stronger conformist transmission biases than those who scored in the middle. In Experiment 1, which is arguably more sensitive than Experiment 2 because there are often more than 2 options, conformist biases strengthen among older individuals and weakens among males.

	Experiment 1	Experiment 2
(Intercept)	-0.31 (-0.76,0.14)	0.33 (-1.64,2.29)
zIQ	0.09 (-0.16,0.33)	0.09 (-0.18,0.37)
zIQ ²	0.26 (0.11,0.41)***	0.25 (0.09,0.42)**
zPrestige	0.04 (-0.18,0.26)	-0.10 (-0.45,0.26)
East Asian	0.36 (-0.15,0.87)	-0.20 (-0.77,0.36)
Other Ethnicity	0.49 (-0.11,1.10)	0.07 (-0.60,0.75)
zAge	0.39 (0.17,0.60)***	-0.05 (-0.29,0.20)
Male	-0.51 (-0.93,0.09)*	0.04 (-0.43,0.51)
<i>R</i> ²	0.33***	0.17 ⁺
N	74	74

***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$

Given the effect of IQ on the amount of social learning and the strength of the conformist transmission bias, a reasonable question is whether these individual differences result in differences in performance and therefore payoffs. A regression analysis of performance on individual predictors revealed a consistent, but weak and non-significant positive effect of IQ on performance (both before and after seeing social information), suggesting that if IQ is helpful in this task, the effect is very weak (see Appendix B.5 for details).

4.7 Discussion

Across two experiments and an ethnically diverse sample, we tested the effect of number of options, transmission fidelity, and payoff size on the degree of social learning and the strength of the conformist bias. Our major findings can be summarized as follows:

Substantial conformist transmission. In both experiments, we found substantial reliance on conformist biased social learning, with only 3% and 9% (or 15%) showing no conformist biases in Experiments 1 and 2, respectively. We suspect the stronger biases in Experiment 1 resulted from having multiple options at play. Some past experiments suggested no conformist bias (Claidière, Bowler, Brookes, Brown, & Whiten, 2014; Claidière, Bowler, & Whiten, 2012; Coultas, 2004; Eriksson & Coultas, 2009). These studies differed from our results in at least two critical ways making them difficult to compare. First, they did not incentivize performance or have a “right” answer. The models tested here make predictions about traits with fitness consequences. In the real world, these may be direct – e.g. eating the wrong kind of berries can get you killed – or indirect – e.g. a norm to cooperate can increase the fitness of the cooperative group. Second, conformity was operationalized with an assumption that conformist transmission requires a neutral prior. Our results show both the presence of a prior and substantial conformist transmission.

Although not directly applicable to the current experiment, there are a few other studies that are worth mentioning in the broader context. Kameda and Nakanishi (2002) show that the conformist bias weakens as the cost of individual learning increases due to a producer-scrounger dynamic. This dynamic may not be applicable in an environment in which information is individually costly to acquire, but cheap to acquire via social learning (as is the often the case in our world of accumulated culture), restoring the adaptive value of a conformist bias. Efferson et al. (2007) field-based transmission experiment is more puzzling. Although a clear signal of majority behaviour was present, many individuals did not conform to this behaviour. These results may indicate cross-cultural variation in the tendency to conform (as has been shown in social learning more generally, see Mesoudi et al., 2015, but not in this experiment), but as the authors discuss, this experiment may be capturing other variables that have yet to be theoretically account.

Increased social learning and stronger conformist bias as the number of options increases. Both the amount of social learning and the strength of conformist biases increased as the number of options increased, as illustrated in Figure 4.3 and Figure 4.4. The increase in social learning corresponds to a “copy when uncertain” strategy; uncertainty increases with number of traits. Together, these results mean that all prior experiments have merely established a lower bound on the amount of social learning and strength of conformist transmission, since all use only 2 options.

Changing inflection point with more options. The inflection point for conformist transmission behaves in a pattern consistent with the theory developed in Nakahashi et al. (2012), except that it is substantially and consistently upward biased. We suspect that this is due to a lack of any account of people’s asocial priors in the Nakahashi et al. model. Future models should include asocial priors.

More reliance on social learning, but stable conformist bias across different transmission fidelities. Unexpectedly, except at very low transmission fidelities (40% error), the strength of conformist transmission was relatively stable and flat across a wide range of transmission fidelities. Though not formally modelled, this pattern seems inconsistent with what we inferred by considering Henrich and Boyd (2002) together with Nakahashi et al. (2012). Three different factors may be relevant. First, the spatial variation typically modelled may be different from transmission errors in some fundamental way, leading us to make an inferential mistake. A proper model of transmission error is required. Another possible issue is that these results are constrained by the limited degrees of freedom in our experiment. That is, in theoretical models (and the real world) where many different types of errors can be made, conformist transmission is adaptive when transmission fidelity is low as these mistakes may result in small improvements. However, by constraining our experiment to two options, of which only one is correct, mistakes are always fatal (win-lose). New experimental designs and more data are needed to address this discrepancy. Finally, it could simply be that human psychological mechanisms are not designed to intuitively evaluate the format in which we provided the transmission fidelities – probabilities of accurate social information – a wealth of research suggests that people are bad at using probabilities (Tversky & Kahneman, 1981). But, since we do observe some effects on social learning, this can't be the complete explanation.

Higher payoffs have little or no effect on learning strategies. The amount of social learning differs between no payoff and some payoff, but does not continue to increase with higher payoffs (Table 4.2). The strength of conformist transmission increases as the payoffs for correct answers increase. This result is not significant (Figures 4.6b and 4.6c), however, Nakahashi et al. (2012) predict a very small effect, so it may be that our transmission error and payoff range were too

small to detect the pattern (see Mathematica file available with the forthcoming paper). Note that here, payoffs relate to the task itself and not the payoff of each individual as in past experiments (e.g. Mesoudi, 2011). Future work might also explore the effect of different fitness landscapes on conformist biased social learning.

Group size affects social learning with 2 options. Consistent with King and Cowlshaw (2007) and Perreault, Moya, and Boyd's (2012) theories, we find that increased group size predicts increased social learning independent of the frequencies of options. However, we did not find this relationship for more than two options. One possibility is that with increased traits, larger groups are required for group size to have a discernible effect (our range of group sizes was 5 to 11).

Cognitive ability differences are associated with both social learning and the strength of the conformist bias. Extrapolating from the existing modelling work, we suspected that IQ would be negatively related to social learning and the strength of the conformist bias. This is the case for social learning, but only the case for the conformist bias in the lower range of IQs. At the upper end, higher IQs, like very low IQs, are associated with stronger conformist biases. These results together suggest that higher IQ individuals are strategically using social learning (using it less, but with a stronger conformist bias when they choose to use other information). However, IQ is only weakly related to overall performance, suggesting that, even if this is the case, these strategies are not particularly effective. Assuming our results generalize to other tasks, differences in cognitive ability may also help explain individual variation in social learning and conformist transmission in non-human species (Laland, Atton, & Webster, 2011; Pike & Laland, 2010).

No detectable 'cultural' differences. Neither our East Asian ethnicity variables nor our cultural identification or acculturation index pointed to any variation in social learning or conformist transmission across these populations. Nevertheless, although 53% of our sample was East Asian

and 85% of them were born outside of Canada, we should take this as only preliminary evidence. It would be preferable to measure East Asians living in East Asia rather than rely on acculturation or cultural identification measures to compensate for the partial acculturation of our mostly WEIRD Canadian sample.

No detectable relationship between prestige and social learning. We predicted that individuals who view themselves as prestigious compared to others may be disinclined to copy others, because they don't see others as superior sources of information. However, we found no relationship between our measure of self-reported prestige and social learning. One reason for this might be that this general sense of prestige is psychologically very distant from the skill domain of line-length judgments, since line-length judging is not a valued skill in Vancouver. Thus, broadly prestigious individuals may not have mapped this over to the experimental task. Further research on this requires using tasks involving locally esteemed skills.

No detectable relationships between other individual variables and social learning or the strength of the conformist bias. Our measures of dominance, rule-following, reflective thinking, or any of the Big 5 personality dimensions did not reliably predict social learning nor the strength of the conformist bias. Thus, our results suggest that conformist biases are not a feature of personality, or other dispositional or normative tendencies like rule-following. Finally, though we were able to account for between 9% and 33% of the variance in individual's reliance on social learning and strength of conformist biases, there remains an immense amount of individual variation in these strategies that we could not explain.

Overall, our findings support the value of formal evolutionary modelling in developing and testing theories about human psychology and about social learning in particular. Broadly, they

indicate that at least in this domain conformist transmission is a central component of human social learning, which varies predictably across contexts and individuals.

Chapter 5: Cultural Dispositions, Social Networks, and the Dynamics of Social Influence

Attitudes and personality traits differ, not only across individuals, but also across entire populations. Compared to North Americans, for instance, people in India and China are generally less extraverted and also endorse less individualistic values (Hofstede, 2003; McCrae, Terracciano, & 79 Members of the Personality Profiles of Cultures Project, 2005). These differences have implications for behavioral outcomes (e.g. extraverts have more acquaintances and individualists are less likely to conform to majority opinion (Asendorpf & Wilpers, 1998; Pollet, Roberts, & Dunbar, 2011)). The psychological study of cultural differences typically focuses on these kinds of individual-level outcomes.

Individual level behavioral outcomes can have further consequences that transcend a psychological level of analysis—emergent consequences that, over time, play out across entire populations (Kameda, Takezawa, & Hastie, 2003; Kenrick, Li, & Butner, 2003; Latané, 1996; Mason, Conrey, & Smith, 2007; Oishi, 2014; Smaldino, 2013; Talhelm et al., 2014; Vallacher, Read, & Nowak, 2002). These consequences are of interest not only to psychologists but also to other scholars who study the things that define societies and populations (public opinion, political ideologies, religious beliefs, etc.), and the speed with which those things change over time. The purpose of this article is to identify the effects that cultural differences in basic behavioral dispositions may plausibly have on these kinds of long-term societal outcomes, and to do so in an analytically rigorous manner.

The analytical method we employ is computational modeling. We report outcomes compiled from tens of thousands of computer simulations, each of which simulated tens of thousands of interactions between individuals within a population. Our models were informed by results of

previous empirical research on social interaction and social influence—including results documenting population-level differences in extraversion and conformity. These models reveal predictable implications of these differences for the emergent properties of the social networks that govern interpersonal interactions, and further implications for the societal outcomes of interpersonal influence within those social networks. We focus on two specific kinds of societal outcomes: (a) the consolidation of majority opinion (the extent to which existing opinion majorities become bigger majorities over time); and (b) the diffusion of innovations (the extent to which new opinions, radical beliefs, and other new ideas spread within a population over time). The results of these models therefore reveal specific ways in which population-level differences in psychological traits may have long-term consequences for cultural stability and cultural change.

5.1 Cultural Differences in Extraversion and Conformity

Numerous results reveal cross-cultural differences in personality traits (Heine & Buchtel, 2009)¹¹. A trait of particular relevance here is extraversion. Multiple studies—employing multiple methods to assess the personality traits of tens of thousands of individuals in dozens of countries worldwide—have revealed differences in mean levels of extraversion (McCrae, 2002; McCrae et al., 2005; Schmitt, Allik, McCrae, & Benet-Martínez, 2007). These differences are associated with differences in conceptually relevant behavioral outcomes (Matsumoto, Yoo, & Fontaine, 2008). The magnitude of cross-cultural differences in extraversion is not huge, but nor is it trivial. For example, individuals living in Morocco have mean extraversion scores that are approximately half a standard

¹¹ We refer to these personality differences as cultural differences and there is evidence that these differences are indeed cultural. For example, bilingual people frame-switch and show different “personalities” depending on the language they’re using (Chen & Bond, 2010; Ramírez-Esparza, Gosling, Benet-Martínez, Potter, & Pennebaker, 2006). However, these differences may also be genetic, evoked, or some combination of these. Our model assumes individual stability in these personality traits, but an interesting additional question would be to look at the transmission and adaptive value of different personality traits in different contexts.

deviation lower than the worldwide mean, whereas individuals in Northern Ireland have mean extraversion scores that are approximately half a standard deviation higher than the worldwide mean (McCrae et al., 2005).

Given these differences in extraversion, one would also expect cultural differences in interpersonal behavior and the measurable outcomes of interpersonal behavior. One of the most obvious outcomes associated with extraversion is the formation of social connections with other people. Compared to more introverted individuals, extraverts have more friends and acquaintances (Kalish & Robins, 2006; Pollet et al., 2011) and the social networks of extraverts grow more rapidly over time (Asendorpf & Wilpers, 1998). These individual differences are reflected in population-level differences: In populations characterized by relatively higher levels of extraversion, people generally have larger networks of friends and acquaintances (Chua & Morris, 2006; Harihara, 2014).

While cross-cultural differences in extraversion have implications for the nature of individuals' interpersonal relationships, other cross-cultural differences have implications for social influence within those relationships. Many different dispositional tendencies are relevant to social influence processes, including basic personality traits (such as openness to experience), authoritarian attitudes, and the endorsement of individualistic versus collectivistic values. There are well-documented cultural differences on these constructs (Farnen & Muelen, 2000; Hofstede, 2003; McCrae et al., 2005). These cultural differences have many implications for individual behavior (Gelfand et al., 2011; Heine & Buchtel, 2009).

One implication is of particular relevance here: The tendency to either conform to, or to deviate from, others' attitudes and actions. Lower openness, higher authoritarianism, and more collectivistic values all imply a greater tendency to conform to perceived social norms; whereas higher openness, lower authoritarianism, and more individualistic values all imply an increased

tendency to resist conforming. Much empirical research shows that, in prototypically collectivistic countries—which are also characterized by lower levels of trait openness and greater endorsement of authoritarian attitudes—people more readily conform to perceived majority opinion (Bond & Smith, 1996; Gelfand et al., 2011; Mesoudi et al., 2015). Note, here we are referring to the tendency to copy majorities rather than the tendency to copy majorities at a rate higher than the rate of the majority (conformist transmission) – conformity in the psychological sense and not the cultural evolutionary sense. For an attempt to connect these two “conformities”, see Claidière and Whiten (2012).

The preceding paragraphs identified population-level differences in behavioral outcomes that are typically measured at an individual-level of analysis (the sizes of individuals' friendship networks, and the tendency of individuals to conform to perceived majority opinion). When aggregated across individuals within any population, and also aggregated across multiple opportunities for interpersonal interaction, these individual-level outcomes can have implications that transcend the individual-level of analysis and must be measured at the level of the populations. Given the implications that extraversion has for the size of individuals' friendship networks, cultural differences in extraversion are likely to have further implications for the structural geometry of the social networks that define entire populations. And, given the implications that conformist attitudes have for actual conformity to majority opinion, cultural differences in conformity are likely to have further implications for the societal outcomes of interpersonal influence within these social networks.

5.1.1 Implications for the Structure of Social Networks

How might the mean level of extraversion within a population affect the structure of the population-wide social network? To address that question, it is first necessary to consider the

geometric properties of these networks of interpersonal connections within human populations. Empirical evidence from many different kinds of populations (Apicella et al., 2012; Henrich & Broesch, 2011; Ugander, Karrer, Backstrom, & Marlow, 2011) show that human social networks have several defining structural properties. One property refers to the frequency distribution of the number of acquaintances that people within a population have (in the network sciences, this is often referred to as a “degree distribution”). Within real human populations, most individuals have at least a few acquaintances, but relatively few individuals have an extremely large number of friends. Consequently, human social networks are characterized by a degree distribution skewed to the right. A second property refers to the likelihood that any two acquaintances of any individual will also be acquainted with each other. Within real human populations, this likelihood is non-zero, which is reflected in indices that assess the “clustering” of social connections within the network. A third property refers to the average smallest number of social connections required to trace a path from any one individual within the population to any other individual within the population. (This is sometimes referred to as “average path length” or, in common parlance, “degrees of separation”.) While there is considerable within-population variability in the path length separating any two individuals, in human social networks the mean shortest path length is typically between 3 and 4.

These network properties are emergent consequences of individuals’ behavioral actions—specifically, their tendencies to make acquaintances with other individuals. Cultural differences in extraversion have an obvious implication. In populations characterized by higher mean levels of extraversion, a greater number of people are likely to make a greater number of acquaintances, and this will result in denser social networks (specifically, a less skewed degree distribution, a higher level of clustering, and a lower mean path length). These structural properties may have evolved for more efficient transmission of information (Pasquaretta et al., 2014).

Social connections are the conduits through which socially contagious things spread throughout human populations. The category of contagious things includes not only socially transmitted diseases, but also socially transmitted information of any kind: ideas, technologies, opinions, beliefs, patterns of behavior, and so forth (Berger, 2013; Eubank et al., 2004; Fowler, Christakis, Steptoe, & Roux, 2009; Rogers, 2003). Thus, if cultural differences in extraversion have implications for the structure of social networks, these cultural differences—along with cultural differences in individuals' tendency to either conform or deviate from the majority opinion—may have further implications for the societal outcomes of social influence. It is to these further implications that we now turn.

5.1.2 Implications for Societal Outcomes of Social Influence Processes

Within the psychological sciences, the study of social influence typically focuses on the processes through which individuals are influenced by, or exert influence on, other individuals, and on variables that affect those individual-level outcomes (Cialdini & Goldstein, 2004)). People are neither simply the targets of influence nor simply the sources of influence; they are both. Over time, people have repeated opportunities to be influenced by, and to exert influence on, other people within their social networks. Thus, when considered within the context of whole populations, social influence is a bi-directional dynamic process, and this has consequences for the patterns of belief and behavior that define populations. It is through this dynamic social influence process that fads and fashions wax and wane, that pockets of public opinion propagate across entire populations, and that radical ideas sometimes catch on and sometimes don't (Harton & Bourgeois, 2003; Kashima, Wilson, Lusher, Pearson, & Pearson, 2013; Latané, 1996). We focus here on two specific population-level phenomena that depend on this dynamic process through which people influence each other, and which are themselves of considerable interest within the social sciences.

5.1.2.1 Consolidation of Existing Opinion Majorities

We focus first on the tendency for existing opinion majorities to become bigger over time—the phenomenon that Latané (1996) labeled *consolidation*. Consolidation of majority opinion emerges as a consequence of the individual-level psychology of social influence, whereby people are inclined to conform to the actions, attitudes and opinions that they perceive in the majority of others (MacCoun, 2012). Individuals who already are in agreement with the perceived majority tend to maintain that opinion over time; individuals whose personal opinions are at variance with the perceived majority feel pressure to change and to adopt the majority opinion instead. Thus, in the absence of countervailing pressures, the size of opinion majorities within a population tends to become incrementally greater over time.

This consolidation phenomenon is relevant to many specific outcomes of considerable societal importance. For instance, it has implications for intergroup prejudice. To the extent that a particular prejudice is perceived to be popular, people are more likely to express that prejudice themselves (Crandall & Eshleman, 2003). In doing so, they reify that existing prejudice and perpetuate it within the society. Consolidation also lies at the root of “bandwagon effects” in electoral politics, in which information about others’ voting intentions may cause previously uncommitted voters to adopt the perceived majority opinion (Kenney & Rice, 1994; Nadeau, Cloutier, & Guay, 1993)—with potentially nontrivial consequences for election outcomes.

5.1.2.2 Diffusion and Spread of New Ideas

Second, we focus on the extent to which new ideas, radical beliefs, and novel ways of doing things spread through a population—the phenomenon that sociologists refer to as the *diffusion of innovations* (Rogers, 2003). If consolidation of majority opinion represents a sort of cultural entrenchment, the diffusion of innovations is a hallmark of cultural change. Not all innovations do

spread, of course. Indeed, the conformist social influence processes that underlie consolidation of majority opinion can pose a substantial psychological barrier to the spread of unpopular attitudes and practices (Eriksson et al., 2007). And yet, as psychological research on minority influence reveals, this barrier can be breached (Wood, Lundgren, Ouellette, Busceme, & Blackstone, 1994); and, as the sociological literature reveals, some innovations do diffuse widely throughout entire populations (Rogers, 2003; Wejnert, 2002).

Because there are so many different kinds of “innovations”—new opinions, new beliefs, new technologies, etc.—the process by which innovations spread (or fail to spread) has implications for many different kinds of societal outcomes. Diffusion processes are of substantial relevance to consumer behavior (Berger & Schwartz, 2011; Brown & Reingen, 1987), to the success or failure of public health interventions (Haider & Kreps, 2004), and to the popular ascendance of novel ideologies and religious beliefs (Collar, 2007), among other specific societal implications.

5.1.2.3 Obvious and Non-obvious Effects of Cultural Differences

How might cultural differences in extraversion and conformity affect the dynamic social influence processes that underlie consolidation of existing opinion and also underlie the diffusion of innovations? At the individual-level of analysis, some initial implications are obvious: Within more extraverted populations—characterized by denser social networks—a greater number of individuals have the opportunity to influence, and be influenced by, a greater number of acquaintances. And, within more conformist populations, individuals are more likely to conform to the actions and attitudes expressed by the majority of their acquaintances. But, with perhaps one exception (consolidation of existing opinion majorities is likely to occur more rapidly in more conformist populations), it is difficult to confidently intuit or logically deduce what further effects these individual-level outcomes might have for the speed with which opinion majorities consolidate within

the population, or for the likelihood that radical new ideas might successfully diffuse throughout a population. Indeed, one of the hallmarks of the non-linear nature of dynamical social influence (and of complex dynamical systems more generally) is that emergent population-level outcomes not only defy intuitive appraisal, they also cannot reliably be predicted on the basis of the linear *if-then* rules that govern deductive analysis (Kameda et al., 2003; Latané, 1996; Mason et al., 2007; Vallacher et al., 2002). In order to plausibly identify the implications that cultural differences might have for consolidation and diffusion, it is helpful to employ the powerful analytic tools of computational modeling.

5.2 Computational Modeling of Social Interaction and Social Influence

Theorizing in psychological sciences typically begins with the identification of some set of assumptions and then proceeds to identify further implications that follow logically from those assumptions. In many cases, natural language structures (e.g., words and their accepted meanings) are suitable for this task. For more complicated psychological processes, it may be preferable to translate psychological constructs into mathematical symbols and equations to ensure the necessary analytic rigor. And in some cases, the level of conceptual complexity may transcend the limitations of natural language and analytically solvable equations, in which case a rigorous approach to the problem may require what Ostrom (1988) called the “third symbol system”: computational modeling.

Computational models are especially useful tools for identifying the ways in which processes that unfold over time at one level of analysis might produce emergent properties measurable at another level of analysis. These tools have proven indispensable in the study of evolutionary biology, behavioral ecology, epidemiology, and meteorology (Bower & Bolouri, 2001; Epstein, 2006; Johnson, 2001; Kitano, 2002; Mangel & Clark, 1988), as well as in the study of cognitive and social

psychology (Hastie & Stasser, 2000; Kenrick et al., 2003; Monroe & Read, 2008; Nowak & Latané, 1994; Pfau, Kirley, & Kashima, 2013). Computational models have been extensively employed in the psychological sciences to study group-level and population-level outcomes of interpersonal influence processes (Hastie & Kameda, 2005; MacCoun, 2012; Nowak, Szamrej, & Latané, 1990; Tanford & Penrod, 1983, 1984). For example, in developing dynamic social impact theory, Latané and colleagues (Latané, 1996; Latane & Bourgeois, 2001; Latané, Liu, Nowak, Bonevento, & Zheng, 1995) programmed cellular automata models to simulate human populations governed by a few rudimentary social psychological facts (e.g., people are more likely to communicate with other people who are closer in geographical space; people mutually influence each other during the course of communication). Although conceptually unremarkable at an individual level of analysis, these models produced notable population-level outcomes, some of which were relatively straightforward (consolidation of majority opinion) and others that were more subtle and surprising (over time, previously uncorrelated beliefs and behavioral patterns become correlated). These outcomes represented a set of scientific hypotheses—arrived at rigorously via computational means—that consequently were tested by empirical evidence (Cullum & Harton, 2007; Harton & Bourgeois, 2003; Harton & Bullock, 2007).

Analogously, in order to address our research questions, we too needed to computationally simulate a set of empirical facts evident in the psychological literature—including cultural differences in dispositional tendencies toward extraversion and conformity, the effects of individuals' dispositional tendencies on the forging of acquaintances, as well as the effects of their dispositional tendencies on conformity to perceived majority opinion. And we too needed to measure a set of population-level outcomes produced by these models: (a) structural properties of emergent social networks, (b) consolidation of majority opinion over time, and (c) spread of innovations over time.

5.2.1 Overview of Our Computational Modeling Methods

Cultural differences in extraversion should have implications for the geometric properties of the social networks that emerge within different populations. And cultural differences in conformity should have implications for the emergent consequences of interpersonal influence within those social networks. To address our research questions, our models therefore required two distinct phases. Phase 1 was designed to model the process through which individuals form acquaintances and, as a consequence, social network structure emerges within a population. It was within the context of this phase that we examined how cultural differences in extraversion may have an impact on the emergent structure of social networks. The second phase (Phase 2) built upon the results of the first, and was designed to model the process through which individuals influence, and are influenced by, other individuals to whom they are connected within a social network. It was within the context of Phase 2 that we used additional methods to measure the consolidation of majority opinion and the diffusion of initially unpopular beliefs, and examined how these outcomes may be affected by cultural differences in both extraversion and conformity.

In the following sections, we describe these methods, and the emergent consequences, in detail. We first describe the manner in which our models operationalized both within-culture and between-culture differences in dispositional tendencies toward extraversion and conformity. We then describe Phase 1 of our models (the emergence of social network structure) along with the results that emerged from this first phase. The methods (and results) described in these sections are simply preliminary steps toward the two main parts of our analysis, both of which focus on Phase 2 of our models (during which we model the population-level consequences of interpersonal influence within social networks). In one section, we describe implications for the consolidation of existing majorities. Results of these models reveal that the speed with which small majorities become larger

majorities is likely to be affected not only by cultural differences in conformity, but also by cultural differences in extraversion. In a subsequent section, we describe implications for the diffusion of innovations. These results reveal that the speed with which initially unpopular beliefs spread within a population is likely also to be affected by cultural differences in conformity (and the exact nature of these effects may strike some readers as somewhat surprising).

5.2.2 Simulation of Individual Differences and Cultural Differences

In our models, we created populations comprised of 900 individuals—a size large enough to be plausibly analogous to meaningful populations (e.g., small-scale societies of the sort studied by ethnographers), while not so large as to be computationally intractable. Each individual within a simulated population was assigned a numerical value representing a dispositional tendency toward *extraversion*, and another numerical value representing a dispositional tendency toward *conformity*. Both extraversion and conformity were operationalized as behavioral probabilities. An individual's extraversion value represented the probability that they would make a new acquaintance when given the opportunity. An individual's conformity value represented the probability that they would change a preexisting attitude (or belief or behavioral practice or any other thing that might be responsive to social influence) upon discovering that the majority of their acquaintances had a different opinion (or belief, etc.).

In assigning these values, we attempted to accomplish two objectives. (1) Within any single simulation, the distribution of values should plausibly mimic individual differences in behavioral dispositions that exist within any human population; and (2) across different sets of simulations, these distributions should plausibly mimic differences between different populations (i.e., realistically represent the magnitude of actual cultural differences). To accomplish these objectives, we drew upon the *beta distribution* (Gupta & Nadarajah, 2004), which can be used to model both within-

population and between-population variability (Balding & Nichols, 1995; Batchelder, 1975). The beta distribution is a family of probability frequency distributions, the shapes of which are controlled by two parameters (denoted $[\alpha, \beta]$). By adjusting these parameters, it is possible to create a wide range of realistic distributions that vary in shape and central tendency. Beta distributions are defined over the probability interval $[0, 1]$, which makes them useful for modeling any underlying variable bounded by two known endpoints. It is especially useful in models—such as ours—in which individual differences are operationalized as behavioral probabilities, allowing us to use a beta distribution without any transformation.

In order to assign extraversion values to individuals within our simulated populations, we created 3 different beta distributions with the following parameter values: $[4, 4]$, $[2.5, 3.5]$, and $[3.5, 2.5]$. The first set of parameters creates a bell-shaped distribution that is symmetrical around a mean value at the midpoint of the probability scale. It represents a kind of “baseline” population in which there are an equal number of introverts and extraverts. The second set of parameters creates a distribution that is skewed right (i.e., introverts outnumber extraverts), and has a mean value approximately 0.5 standard deviations less than the baseline population. The third set of parameters creates a distribution that is skewed left (i.e., extraverts outnumber introverts), and has a mean value approximately 0.5 standard deviations higher than the baseline population. (See Figure 5.1 for a graphical representation of the three beta distributions.)

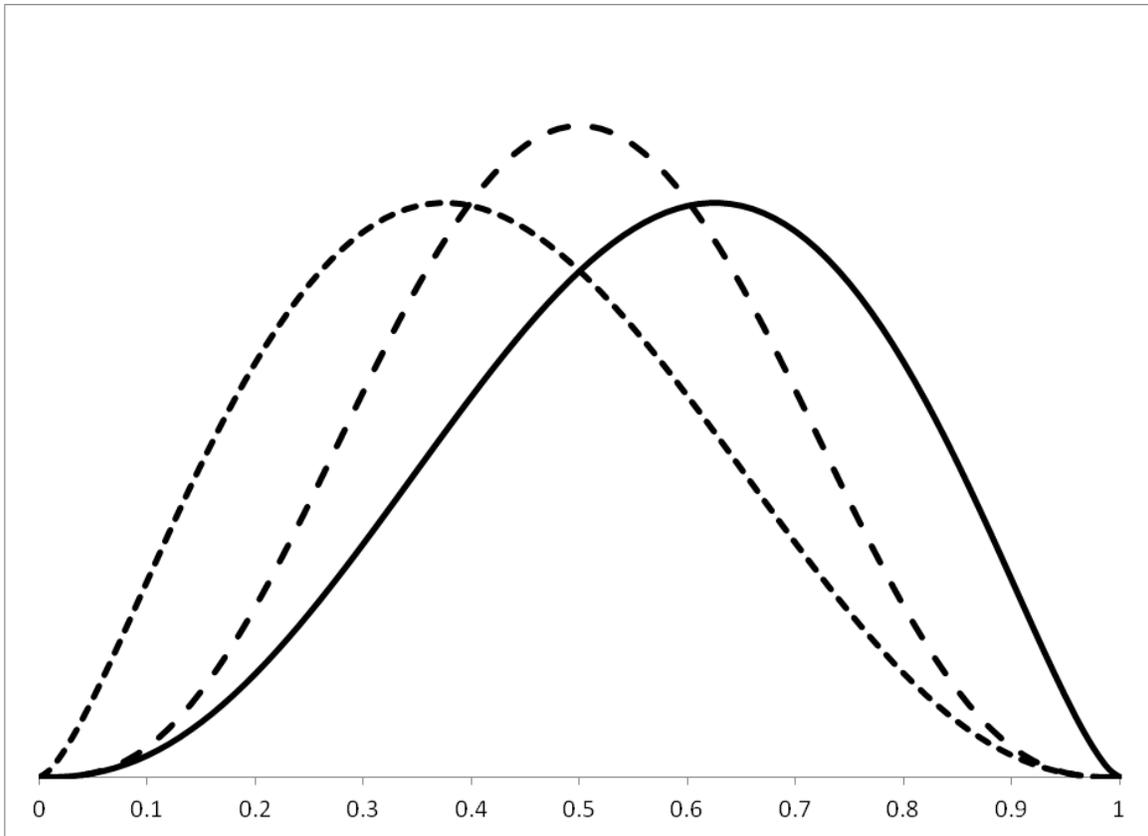


Figure 5.1 Three beta distributions from which values were randomly drawn to simulate individual-level differences and population-level differences in dispositional tendencies toward extraversion and conformity. The symmetrical distribution (long-dashed line) represents individual differences within populations with a moderate mean level of the disposition (equal to the global mean), such as Peru. The right skewed distribution (short-dashed line) represents individual differences within populations with a relatively low mean level of the disposition (approximately 0.5 standard deviations lower than the global mean), such as Morocco. The left skewed distribution (solid line) represents individual differences within populations with a relatively high mean level of the disposition (approximately 0.5 standard deviations higher than the global mean), such as Northern Ireland.

For each simulation, each of the 900 individuals within the population was randomly assigned an extraversion value drawn randomly from one of these three beta distributions. For some

simulations, values were drawn from the $\beta[4,4]$ distribution; consequently, these simulations represent populations with a moderate level of extraversion. For other simulations, values were drawn from the $\beta[2.5, 3.5]$ distribution and represent populations with a relatively low level of extraversion. And for still other simulations, values were drawn from the $\beta[3.5, 2.5]$ distribution, and represent populations with a relatively high level of extraversion. This ensured a realistic representation of individual differences within each simulated population. Also, because differences between the means of the 3 beta distributions mathematically mimic the magnitudes of actual cross-cultural differences in extraversion (McCrae et al., 2005), this procedure also created realistic representations of different populations with either moderate (e.g. Peru), low (e.g. Morocco), or high (e.g. Northern Ireland) mean levels of extraversion. By restricting our model to a realistic range we are able to theorize about plausible implications of differences in extraversion within real human populations rather than theoretically possible effects if extraversion were higher or lower.

We used an identical procedure to also assign each individual a probability value corresponding to a dispositional tendency toward conformity. Thus, within each individual simulation, the procedure simulated individual differences in conformist tendencies; and, across all simulations, the procedure created realistic representations of different populations characterized by either moderate, low, or high mean levels of conformity.

In reality, cultural tendencies toward extraversion and cultural tendencies toward conformity are inversely correlated (Hofstede & McCrae, 2004; Schaller & Murray, 2011) and it would be an interesting and important follow-up to look at the effect of this correlation. By definition, however, these constructs are distinct, and they are likely to have conceptually separable consequences on societal outcomes. Therefore, we assigned Extraversion values and Conformity values independently. Across the full set of simulations, we created 9 conceptually distinct types of

populations by crossing the 3 levels of Extraversion and the 3 levels of Conformity in 3 x 3 factorial design. For each of these 9 types, we employed our sampling methods to create 10 different 900-individual populations, ensuring that the simulation results would not be idiosyncratic to any single population of 900 individuals.

5.2.3 Phase 1: Emergent Differences in the Structure of Social Networks

Following the creation of a population, the first phase of our simulations was designed to model a small set of decision rules that govern the formation of social connections between individuals and thus, over time, lead to the emergence of social network structure within the entire population. Within the network sciences, there exist many computational algorithms that can lead to the emergence of some kind of network structure (Jackson, 2010); but many of these algorithms fail to produce the structural properties of real human *social* networks, or fail to do so in a manner that is behaviorally realistic (Schnettler, 2009). For our purposes, it was necessary that our model generated structurally realistic social networks (i.e., social networks with realistic degree distributions, realistic levels of clustering and realistic mean path lengths), and did so through a process that plausibly mimicked the mechanisms through which human social networks form in the real world (i.e., as an emergent property of individuals' behavioral decisions). Furthermore, in order to examine the implications that cultural differences in extraversion may have on emergent social network structure (and, consequently, on the process through which social influence propagates through a population), it was necessary to model the effect that individual differences in extraversion have on the formation of social connections.

Each simulation began with the 900 individuals located in space on a toroidal grid lattice. Each individual was initially assigned exactly four acquaintances: their four closest “neighbors” on the lattice (i.e., the individuals to their immediate east, west, north, and south). The toroidal

geometry ensured no edge effects by connecting individuals on the northern border to those on the south and those on the western border with those on the east. We then allowed the model to iterate. On each iteration, each individual (i) had a probability (p_i)—varying between 0 and 1—of moving to an adjacent space on the lattice. If two or more individuals occupied the same space on any iteration, they “met” and formed an “acquaintance”. These acquaintances were maintained throughout the rest of the simulation and so, over repeated iterations, individuals had the opportunity to accumulate more and more acquaintances. The formation of acquaintances was computationally constrained in two important ways, both of which are informed by the empirical literature on social interaction:

First, the formation of acquaintances was constrained by proximity. Empirical research shows that individuals are more likely to form acquaintances with other individuals who are closer in geographic space (Festinger, Schachter, & Back, 1950; Harton & Bullock, 2007; Latané et al., 1995). It was important to model this constraint because it contributes to the emergence of realistic social network structure. Our model did so by limiting the movement of individuals: On any given iteration, individuals were allowed only to move to an *adjacent* space on the lattice. Thus, from an initial starting configuration, an individual was more likely to befriend those closer in geographic proximity than those further away.

Second, the probability of forming an acquaintance was constrained by individual differences in extraversion. Empirical research shows that more highly extraverted individuals are more likely to form acquaintances with other individuals (Asendorpf & Wilpers, 1998; Paulhus & Trapnell, 1998; Selfhout et al., 2010). To operationalize this principle, each individual's probability (p_i) of moving to a randomly-chosen adjacent space (and thus potentially forming a new acquaintance) was identical to that individual's extraversion value (drawn from the beta distribution; see above). These p_i values remained constant across iterations, thus mimicking the effects that chronic individual differences in

extraversion have on the likelihood of forming new acquaintances. In sum, the algorithm represents a random walk over a grid lattice where the probability of taking a step in one of four cardinal directions—and thus potentially forming a new acquaintance—is given by an individual’s level of extraversion¹².

Social network structure emerges as the model iterates; and as it iterates further—and individuals within population meet more new acquaintances—the social network structure becomes denser. (As the number of iterations approaches infinity, the algorithm generates a network where everyone is directly connected to everyone else.) Given the objectives of this phase of the model, it was necessary to impose a “stopping rule” before the network structure became unrealistically dense. In order to meaningfully compare emergent network structures across different populations (characterized by either low, moderate, or high mean levels of extraversion) that stopping rule had to be identical for every simulation. The stopping rule we chose was simple: We stopped Phase 1 of each simulation after 50 iterations. This stopping rule was informed by the results of preliminary exploratory simulations. These results revealed that—regardless of the mean level of extraversion within a simulated population—50 iterations was sufficient for the emerging social network to attain

¹² Mathematically, the individual’s position (z) after N iterations is commonly expressed in phasor notation (i.e. a complex number as an exponent, with coordinates corresponding to the real and imaginary term):

$$z = \sum_{j=1}^N e^{i\theta_j}$$

Where the angle θ is restricted to one of four cardinal directions $\left(0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\right)$ on a 2 dimensional complex plane representing the 2 dimensions of the grid lattice. (A complex number is a useful way of representing the 2D space, since it has two components—the real portion and the imaginary portion.)

structural properties (degree distribution skew, clustering, path length) that lay within the realistic range of the structural properties that characterize real human social networks. Note, that we are not suggesting that real human populations acquire their characteristic properties after a certain number of interactions – real world interactions continue over time and are subject to generational birth-death processes. Instead, our model offers a convenient method for exploring the effect of population-level differences in personality for population-level differences in social network structures. To show that the properties of these social network structures represent equilibrium differences between populations requires us to incorporate evolutionary birth-death processes into our social network model, which we intend to do in the future.

Recall that for each of the 9 types of populations we created (see above), we created 10 distinct populations. Of these 90 total populations, 30 represented populations with low, moderate, or high mean levels of extraversion, respectively.

The key question addressed in this phase of the model was this: Did the mean level of extraversion within a simulated population influence the structural properties of the social networks that emerged within that population? The answer is provided by results are presented in Table 5.1, which—for each level of extraversion—summarizes mean values for the 3 defining properties of social networks (degree distribution skew, clustering, path length). There are two important aspects to these results. First, the mean values are comparable to values obtained from empirical measurements of the network structure of real populations (Apicella et al., 2012; Henrich & Broesch, 2011; Ugander et al., 2011). This provides reassurance that our modelling methods did lead to the emergence of realistic network structures across all simulated populations. Second, these results reveal population-level differences in the density of the social networks that emerged in the different sets of simulated populations.

We analyzed the difference between the properties of networks that emerge from different levels of extraversion. Within simulated populations with relatively higher mean levels of extraversion, the emergent social networks were characterized by less skewed degree distributions, higher levels of clustering, and lower mean path length. Treating each individual simulation as the unit of analysis, a multiple regression model reveals each of these effects to be statistically significant (p 's < .001).

Table 5.1 Structural properties of the social networks that emerged in Phase 1 of the simulations, as a function of the population-wide mean level of extraversion within the population. Tabled values are means computed across 100 simulations for each of the three levels of extraversion (standard deviations around these means are in parentheses).

Population-Wide Level of Extraversion	Characteristic Path Length	Clustering Coefficient	Degree Distribution Skew
Low	3.82 (.04)	.13 (.005)	.56 (.07)
Medium	3.49 (.02)	.15 (.004)	.37 (.07)
High	3.23 (.02)	.16 (.003)	.25 (.08)

In sum, the first phase of our model produced emergent network structures that closely mimicked the structures of real social networks with real human populations, and the structural properties of those emergent social networks were influenced by cultural differences in extraversion. These results are consistent with empirical evidence documenting cultural differences in social network properties (Chua & Morris, 2006; Harihara, 2014), which further bolsters confidence in the verisimilitude of our computational model. Furthermore, conceptually, these results represent a

means through which population-level differences in extraversion may have further implications for the population-level consequences of interpersonal influence.

5.2.4 Phase 2: Interpersonal Influence within Social Networks

The social network structures that emerged during the first phase of the simulation were kept intact (i.e., we did not allow the structure of those networks to change any further) throughout the second phase—in which we modeled the effects of interpersonal influence within the social networks that characterize different populations. Specifically, we modeled the process whereby (a) individuals obtain information about the opinions and beliefs of their acquaintances, and potentially (b) update their own opinions and beliefs accordingly (depending upon the extent to which their acquaintances' opinions differ from their own, and depending also upon their own dispositional tendency toward conformity). Our methods were designed to realistically model the potential consequences that individual and cultural differences in extraversion may have on social influence processes: Because more extraverted individuals accumulate more acquaintances (as documented in Phase 1), more extraverted individuals also sample the opinions and beliefs of a greater number of other people. Our methods were also designed to realistically model the potential consequences that individual and cultural differences in conformity have on the outcomes of social influence processes: Individuals who are more chronically disposed toward conformity have a higher likelihood of adopting the opinions and beliefs that they perceive to be held by the majority of their acquaintances.

We initiated the second phase of each simulation by assigning one of two possible opinions to each of the 900 individuals within the population. These opinions were binary (0 or 1), and so could conceptually represent any opinion, belief, or behavioral tendency that might be subject to social influence. To ensure that our results were not idiosyncratic to the particular initial assignment

of opinions, we ran 10 different starting positions for each of the 90 populations we created. The specific rules for assigning opinions to individuals differed depending upon whether the simulations were designed to model consolidation of majority opinion or to model diffusion of innovation. (We provide additional details on these assignment rules below.)

We then allowed the model to iterate. On each iteration a single individual was randomly selected to be a target of social influence and so it required 900 iterations for each individual to have, on average, one opportunity to be the target of influence. For the sake of exposition, we may consider every set of 900 iterations to represent one opportunity for influence.

Being the target of influence meant two things: The individual sampled the opinions of their acquaintances in order to determine majority opinion, and then the individual had a probability—varying between 0 and 1—of adopting that majority opinion as well. The sampling of other individuals' opinions was computationally constrained so as to mimic the empirical finding that individuals are influenced not so much by global majorities but by *local* majorities—the opinions that are most popular among the individuals they actually interact with (Cullum & Harton, 2007; Kashima et al., 2013). We modeled this as the majority opinion among the set of acquaintances that the individual had acquired during Phase 1 of the model (see above). The probability that an individual would actually adopt the perceived majority opinion was a joint product of (a) the size of the majority (individuals were more likely to conform to the local majority as the size of that majority increased), and (b) individual's dispositional tendency toward conformity (the value drawn from the beta distribution; see above). The latter values—conformity values—remained constant

across iterations, thus mimicking the effects that chronic individual differences in conformity have on the likelihood that individuals will adopt the opinions of the majority of their acquaintances¹³.

Using these methods, we operationalized individual differences in both extraversion and conformity: Individuals with higher extraversion values were likely to have more acquaintances' opinions to sample when computing the majority opinion; and individuals with higher conformity values were more likely to actually adopt that majority opinion. These individual differences also manifested as cultural differences: In populations with higher mean levels of extraversion, individuals' opinions were (on average) influenced by a greater number of acquaintances' opinions; and in populations with higher mean levels of conformity, individuals were (on average) more likely to adopt the majority opinion expressed by their acquaintances. In the following two sections, we describe in detail the implications that these cultural differences had on the tendency for population-wide opinion majorities to grow larger over time, and on the long-term prospects for new (and initially unpopular) opinions to spread more widely within populations.

5.3 Simulated Effects of Cultural Differences on Consolidation

What implications might cultural differences in extraversion and conformity have for the consolidation of majority opinion over time? To address this question, we ran a total of 900

¹³ Mathematically, the probability (P_{ij}) of an individual (i) acquiring the majority opinion j is given by the following function, where b_j is the number of acquaintances of i with opinion j :

$$P_{ij} = c_i \frac{b_j}{b_0 + b_1}$$

Where c_i represents individuals' conformity value (drawn from the beta distribution, with a value lying within a range from 0 to 1), b_0 is the number of acquaintances who hold opinion 0, and b_1 is the number of acquaintances who hold opinion 1, and b_j represents whichever of those latter two numbers (b_0 or b_1) is greater.

simulations (100 simulations for each of the 9 different populations created by crossing 3 levels of extraversion and three levels of conformity). We initialized each simulation by randomly assigning one of two opinions to each of the 900 individuals within the population. Given that assignment was random, it was very rare that each opinion was held by exactly 50% of individuals. Instead, each simulation began with one of the two opinions being held by a very small majority (typically between 50% and 55% of the total population). As the model began to iterate—and individuals had the opportunity to be influenced by their acquaintances—initial small majorities did not always endure. Regardless, as the model continued to iterate, one of two opinions eventually not only endured as the majority, but also became an increasingly larger majority. The key question here is whether—across all 900 simulations—the speed of this consolidation phenomenon differed across different populations.

There are several complementary analytic approaches that can address that question. One approach is to choose some threshold for the size of a “super-majority,” to measure how many opportunities for influence transpired before a super-majority of that size emerged, and to examine the effects that mean population-wide levels of extraversion (3 levels: low, moderate, high), and mean population-wide levels of conformity (3 levels: low, moderate, high) have on that measure. We conducted analyses for a variety of different super-majority thresholds (e.g., 75%, 90%), and the results were similar regardless of which specific threshold is chosen. We report here the results for a 2/3 super-majority¹⁴.

¹⁴ The 2/3 super-majority corresponds to a decision rule that is commonly used in many real-world decision-making contexts. E.g., in the world's two most populous democracies (India and the United States), constitutional amendments require a 2/3 super-majority vote within the relevant voting bodies.

Figure 5.2 depicts, for each of the 9 populations (100 simulations for each), the mean opportunities for influence required before majority opinion eventually reached the 2/3 super-majority threshold. Two distinct effects can be detected from these results, one of which is more obvious than the other. The obvious effect is a main effect for the mean level of conformity within a population: Opinion majorities more quickly reached the super-majority threshold in populations characterized by relatively higher values of conformity. Less obviously, there appeared also to be a main effect for the mean level of extraversion within a population: Opinion majorities also consolidated into super-majorities more quickly in populations characterized by relatively higher values of extraversion.

These effects are substantiated by the results of a multiple regression analysis that tested the effects of cultural differences in conformity and extraversion on the number of influence opportunities required for the 2/3 supermajority to emerge. These results are reported in Table 5.2. Across all 900 simulations, the main effects of conformity ($p < .001$) and extraversion ($p = .002$) were both statistically significant¹⁵. Drawing on the results of these regression analyses, these effects can be illustrated as follows: Compared to populations with high levels of conformity, low-conformity populations required approximately 0.62 standard deviation more influence opportunities before majority opinion consolidated to the 2/3 super-majority threshold. And, compared to populations with high levels of extraversion, low-extraversion populations required

¹⁵ Results were similar when we conducted analyses that focused on other super-majority thresholds. For thresholds of 75% and 90%, the main effect of conformity was associated with standardized effect sizes of -0.35 and -0.32 (both p 's $< .001$); and the main effect of extraversion was associated with standardized effect sizes of -0.07 and -0.06, (p 's were .062 and .158, respectively.)

approximately 0.22 standard deviation more influence opportunities before the super-majority threshold was reached.

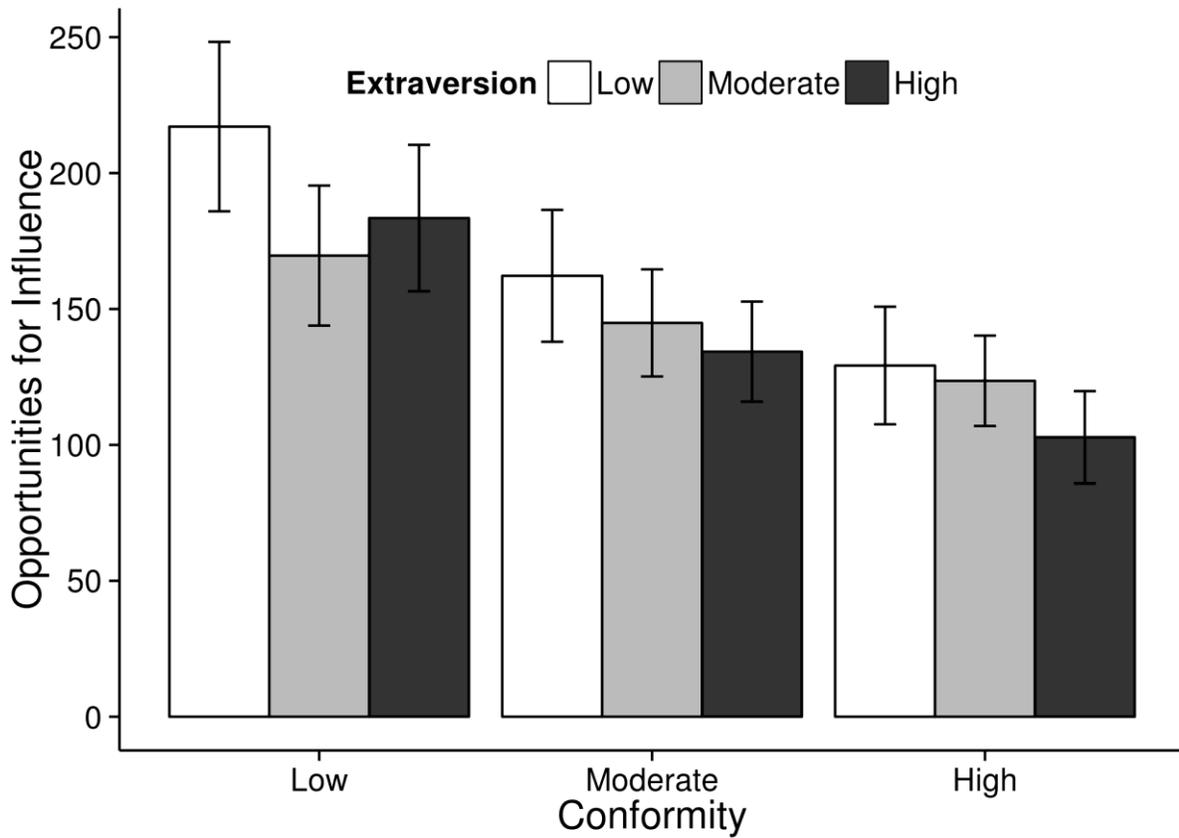


Figure 5.2 Mean number of opportunities for influence elapsed before majority opinion within a population reached a 2/3 super-majority threshold. (Means computed from 100 simulations for each of the 9 cultural populations. Error bars represent 95% confidence intervals).

Table 5.2 Results of multiple regression analysis with random effects for each network, testing the effects that population-wide mean levels of extraversion and conformity had on the log of the number of influence opportunities that elapsed before majority opinion reached a 2/3 super-majority threshold. We used the log of number of influence opportunities due to positive skew in the residuals. R^2 calculated on full model, including random effects (Nakagawa & Schielzeth, 2013).

	<i>B</i>	SE	95% CI	<i>P</i>
Extraversion	-.12	0.04	[-0.19, -0.04]	.004
Conformity	-.31	0.04	[-0.39, -0.23]	<.001
Extraversion x Conformity	.01	0.05	[-0.08, 0.11]	.802
Intercept		0.03	[-0.06, 0.06]	1.00

Note. $R^2 = .07$

Although the means presented in Figure 5.2 offer some hint of an interaction between extraversion and conformity (effects of cultural differences in conformity appeared to be especially pronounced in populations with relatively low mean levels of extraversion), the inferential statistical results (Table 5.2) provide no substantiation for that apparent interaction. It is worth noting that the statistical power of this analysis is constrained by the number of simulations we conducted.

5.3.1 Summary and Discussion

These simulation results show how cultural differences in extraversion and conformity may have implications not only for individual-level outcomes, but also for the population-level phenomenon in which existing opinion majorities become larger over time. The main effect of conformity is unsurprising and its explanation is straightforward: Given that the phenomenon

itself—consolidation of majority opinion—is dependent upon individuals’ tendency to conform to majority opinion, it follows logically that consolidation will occur more rapidly within populations containing a higher number of conformists. The main effect of extraversion is less obvious. In order to make sense of it, it is useful to refer to previous work on the population-level consequences of interpersonal social influence processes. Research on dynamic social impact theory shows that even as opinion majorities grow bigger over time, there still persist subpopulations of people holding the minority opinion (Harton & Bullock, 2007; Latané, 1996). These clusters of unpopular opinion persist in part because the people who comprise those clusters interact primarily with each other, and so are less susceptible to influence by the broader population of people who hold the global majority opinion. In populations with low levels of extraversion, many people are likely to have such circumscribed networks of acquaintances. But, as the mean level of extraversion within a population increases, the number of people who fit this profile decreases. Instead, as extraversion increases, there is also an increase in the percentage of people for whom the local majority (i.e., the majority opinion expressed within one’s personal network of acquaintances) is more diagnostic of the global majority; and so, by conforming to the local majority, they conform also to the global majority, with the consequence that the global majority consolidates more quickly.

Extraversion and conformity both increase the speed of consolidation. However, some readers may find it surprising that the overall model fit is small – 93% of the variance is unexplained (or at least, not attributable to mean levels of extraversion and conformity). However, this result is consistent with other network science research, which suggest that small differences in initial conditions have a large unpredictable effect on information cascades (Watts, 2002; Watts & Dodds, 2007). Despite this large role of randomness, our model suggests that extraversion and conformity both play an important part in consolidation.

5.4 Simulated Effects of Cultural Differences on the Diffusion of Innovations

Although consolidation of majority opinion is defined by some incremental change in popular opinion, it also represents a form of cultural stability—or at least a sort of cultural resistance to the spread of novel or unpopular beliefs. Does this mean that novel and unpopular beliefs are always doomed to failure? Clearly not. Despite their numerical disadvantage, some initially unpopular beliefs do successfully spread within human populations—especially when initial adherents have unshakeable faith in those beliefs and have the motivation and means to influence others (Moscovici, 1980; Wood et al., 1994).

How might the spread of initially unpopular beliefs differ, depending on the mean levels of extraversion and conformity within a population? Intuitively, one might assume that, if consolidation of majority opinion is facilitated by higher levels of conformity and extraversion (as we have just seen), then initially unpopular beliefs are most likely to spread widely in populations characterized by low levels of both conformity and extraversion. As we shall show, computer simulations reveal this intuition to be wrong (as is often the case for intuitions about the outcomes of non-linear dynamical systems). We conducted two sets of simulations, each of which examined diffusion outcomes that resulted from somewhat distinct starting conditions. One set of simulations examined outcomes within a "lone ideologue" context: A situation in which, initially, there is just a single individual espousing an unpopular belief (and doing so with unshakeable faith). The second set of simulations examined diffusion outcomes within a context in which the ideologue is accompanied by a small band of "disciples" who also share the initially unpopular belief (but not their ideological acquaintance's unshakeable faith).

5.4.1 The "Lone Ideologue" Context

We ran a total of 9000 simulations. Specifically, for each of the 9 different kinds of populations—created by crossing 3 levels of extraversion and 3 levels of conformity—we created 100 separate populations within which we simulated the spread of an initially popular belief 10 times each. We initialized each simulation by assigning everyone in the population the same belief, with the exception of 1 individual (i.e., in each simulation, 899 people received one belief and 1 person received a different belief.) In a set of pilot simulations, we discovered that we could not simply choose this 1 lone individual randomly; if we did so, the likelihood of spreading the initially unpopular belief approached zero. Therefore, we did two things to boost the chances that the initially unpopular belief might spread to others. First, we assigned the unpopular belief to the individual within each population who had the highest extraversion value (drawn from the relevant beta distribution, as described above). Second, we re-assigned this individual a conformity value of 0. By taking these two steps, we ensured that this individual had the means to potentially influence many others (because, as a consequence of an unusually high extraversion value, this individual had acquired an unusually large network of acquaintances in Phase 1 of the model), and that this individual was resistant to any pressure to conform to the beliefs expressed by others (all of whom initially held a different belief). As the model iterated—and individuals had the opportunity to be influenced by their acquaintances—there was considerable variability across simulations in the extent to which the initially unpopular belief spread from the lone ideologue to others within the population. The key question here is whether—across all 9000 simulations—the success of this diffusion phenomenon differed across different populations.

As with our examination of the consolidation phenomenon (described above), we chose a specific threshold that defines “successful” diffusion, to measure the percentage of simulations that

eventually reached that threshold, and to examine the effects that mean population-wide levels of extraversion (3 levels: low, moderate, high), and mean population-wide levels of conformity (3 levels: low, moderate, high) had on that measure. In taking this approach, we defined successful diffusion as 50% penetration—the point at which an unpopular belief is transformed into a popular one. Therefore, we examined the effects that cultural differences in conformity and extraversion had on the likelihood that an initially unpopular belief eventually reached this crucial threshold of popular penetration.

Figure 5.3 depicts, for each of the 9 populations, the percentage of simulations (out of a total of 1000 simulations per population) that reached this 50% threshold. Table 5.3 summarizes the results of a binary logistic regression analysis that statistically tests the effects of cultural differences in conformity and extraversion, and their interaction, on the likelihood of reaching this threshold. These results reveal main effects of both conformity and extraversion.

Interestingly (and perhaps contrary to intuition), the effect of conformity was *positive*. An unpopular belief (held initially by just a single well-connected and highly-committed individual) was more likely to successfully spread in populations characterized by *higher* mean levels of conformity. In low-conformity populations, the likelihood was 25% that the initially unpopular belief eventually reached the 50% threshold; but in high-conformity populations, this likelihood increased to 45%.

The results also revealed a *negative* effect of extraversion. An initially unpopular belief was more likely to spread in populations characterized by *low* levels of extraversion. In low-extraversion populations, the likelihood was 40% that the initially unpopular belief eventually reached the 50% threshold; but in high-extraversion populations, this likelihood decreased to 30%.

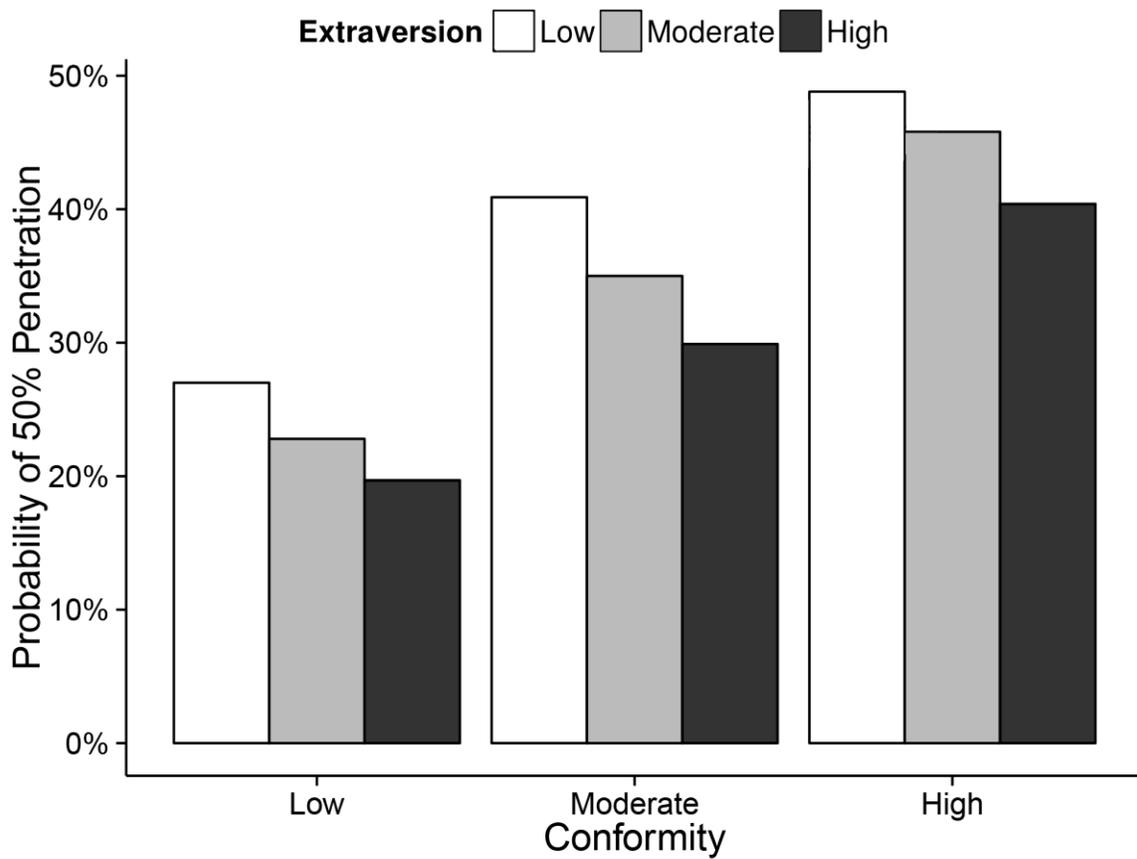


Figure 5.3 Percent of simulations in which a new belief—held initially by just one highly extraverted “lone ideologue”—successfully spread to 50% of the entire population. (Percentage values based on 1000 simulations for each the 9 cultural populations.)

Table 5.3 Results of binary logistic regression analysis with random effects for each network, testing the effects that population-wide mean levels of extraversion and conformity had on the likelihood that a new belief—held initially by just one highly extraverted “lone ideologue”— spread to 50% of the entire population. Pseudo- R^2 calculated on full model, including random effects (Nakagawa & Schielzeth, 2013).

	<i>Odds Ratio</i>	<i>b</i>	SE	95% CI (Odds Ratio)	<i>p</i>
Extraversion	.75	-.29	.14	[.56, 1.00]	.048
Conformity	1.86	.62	.15	[1.39, 2.52]	<.001
Extraversion x Conformity	1.06	.06	.18	[.74, 1.52]	.756
Intercept	.44	-.81	.12	[.35, .56]	<.001

Note. Pseudo $R^2 = .23$

These results reveal a positive main effect for conformity and a negative main effect for extraversion, both of which are consistent with the effects that emerged in the previous analysis: Initially unpopular beliefs spread more readily in populations characterized by higher levels of conformity, and by lower levels of extraversion. Before discussing any of these effects further, it is instructive to examine the extent to which they emerge also under simulated circumstances in which the initially unpopular belief is held not simply by a lone ideologue, but by an ideologue accompanied by one or more disciples.

5.4.2 The "Ideologue Accompanied by Disciples" Context

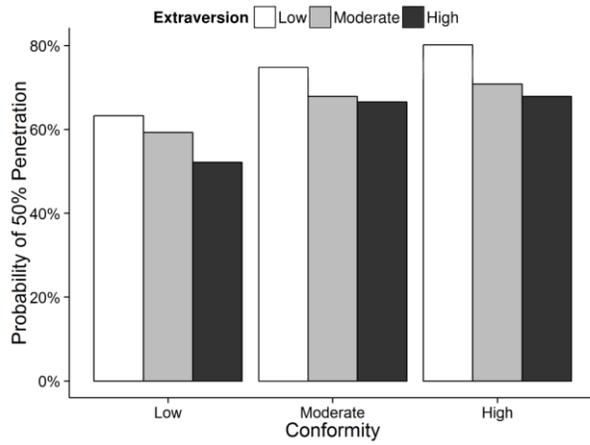
We ran an additional 108,000 simulations to examine a diffusion context in which, rather than being the only initial adherent to an unpopular belief, the ideologue is accompanied by a small

set of disciples who also initially hold the same unpopular belief. Rather than arbitrarily choosing a specific number of disciples, we ran separate sets of simulations corresponding to circumstances in which the ideologue was accompanied by different numbers of disciples, ranging from a single disciple to 12 disciples.

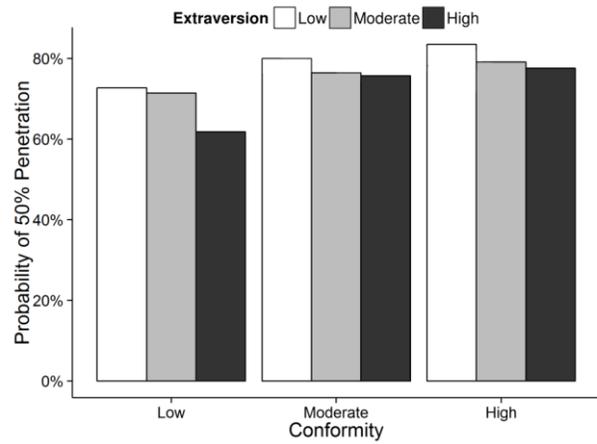
We did so as follows: First, just as in the “lone ideologue” simulations, we assigned the unpopular belief to the individual with the highest extraversion value (and also assigned this individual a conformity value of 0). We then assigned the same belief to 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, or 12 randomly chosen acquaintances (“disciples”) of that individual. These disciples' conformity values were unchanged from the values initially drawn from the relevant beta distribution. (Therefore, while they shared the ideologue's unpopular belief, the disciples do not share the ideologue's unshakeable faith in that belief.) The remaining 887 – 898 individuals (the exact number varied, depending on the number of disciples that the ideologue was assigned) were assigned the opposite belief. We ran 9000 simulations (1000 simulations for each of the 9 different populations) for each of the 12 conditions defined by specific numbers of disciples.

The model then proceeded to iterate. We examined the extent to which the initially unpopular belief spread through the population, and the extent to which diffusion differed across different populations. We did so by focusing on the likelihood that the initially unpopular opinion eventually spread successfully to 50% of the population. Table 5.4 summarizes the key results from 12 separate binary logistic regression analyses, each of which analyzed results from a subset of 9000 simulations associated with a specific number of disciples (ranging from 1 to 12). For each of these 12 regression analyses, odd ratios reveal the effects of extraversion and conformity (and their interaction) on the likelihood that the 50% threshold was attained.

These results reveal three things. First, conformity exerted a positive effect (indicated by odds ratios > 1) and this positive effect emerged regardless of the number of disciples. Second, extraversion exerted a weaker negative effect (odds ratios < 1) up to about 6 disciples. With more disciples, this effect started, was more erratic and sometimes positive, as with the effect of consolidation (as shown in the line graph at the bottom of Table 5.4). Third, there was no consistent interaction between mean extraversion and conformity. (In 9 of the 12 numbers-of-disciples conditions, the odds ratio associated with the interaction was greater than 1; in the other 3 conditions the odds ratio was less than 1. This pattern provides no basis for any confident inference about an interaction.) Finally, the probability of penetration of the new idea increased with number of disciples, as shown by the Intercept. The two main effects (and the lack of a meaningful interaction) are graphically illustrated in Figure 5.4, which summarizes results obtained within the subsets of simulations in which the ideologue was accompanied by either 6 disciples or by 12 disciples. The initially unpopular belief was more likely to successfully spread to 50% of the population within populations characterized by higher levels of conformity and lower levels of extraversion.



6 disciples



12 disciples

Figure 5.4 Percent of simulations in which a new belief—held initially by a highly extraverted ideologue along with either 6 disciples or 12 disciples—successfully spread to 50% of the entire population. (Percentage values based on 1000 simulations for each the 9 cultural populations.)

Table 5.4 Results (odds ratios) of binary logistic regression analyses with random effects for each network, testing the effects that population-wide mean levels of extraversion and conformity had on the likelihood that a new belief— held initially by an ideologue along with disciples——spread to 50% of the entire population. Each row presents results associated with the subset of 9000 simulations associated with a specific number of disciples (varying from 1 to 12).

Number of Disciples	<i>Odds Ratio</i>			
	Main Effect of Extraversion	Main Effect of Conformity	Extraversion x Conformity Interaction	Intercept
1	.885	1.529	1.085	.876
2	.918	1.427	1.328	1.043
3	.855	1.814	1.081	1.601
4	.832	1.792	.979	1.876
5	.886	1.739	.899	2.086
6	.767	1.267	1.094	2.649
7	.947	1.694	1.027	2.373
8	.705	1.398	1.156	3.497
9	.649	1.573	1.152	3.528
10	1.086	1.433	1.094	2.269
11	.704	2.010	1.035	3.831
12	1.240	1.187	0.732	3.557






5.4.3 Summary and Discussion

Just as mean levels of conformity and extraversion within simulated populations had implications for the consolidation of existing majority opinions, so too did they have implications for the successful spread of initially unpopular beliefs. The effect of conformity emerged consistently across both the “lone ideologue” and “ideologue accompanied by disciples” contexts, and is perhaps somewhat counter-intuitive: Initially unpopular beliefs spread more successfully in populations characterized by relatively high levels of conformity. Similarly, the effect of extraversion was weaker than that of conformity, but it emerged consistently across both the “lone ideologue” and “ideologue accompanied by disciples” contexts. This effect too is also perhaps a bit counterintuitive: Initially unpopular beliefs spread more successfully in populations characterized by relatively low levels of extraversion.

How is it that radical new beliefs—which were initially unpopular—spread more successfully in populations with relatively high numbers of conformists? The answer lies in the fact that social influence is governed primarily by *local* norms rather than global norms: When individuals conform, they tend to conform to whatever belief is held by the majority of people in their own personal social networks (regardless of whether or not the locally-popular belief is objectively popular in the broader population). This tendency to conform to local norms occurs more readily among individuals who are more dispositionally inclined toward conformity. Consequently, populations characterized by high levels of conformity are more vulnerable to these local social influence outcomes. This principle applies under circumstances in which local norms match global norms (and so accounts for the faster consolidation of majority opinion); and it also applies under the rarer set of circumstances in which local norms deviate from global norms. It is because of the latter effect

that unpopular beliefs can spread simply as a consequence of conformity processes—and spread more rapidly within more highly conformist populations.

The individuals most likely to perceive a globally unpopular belief to be popular among their acquaintances are those who have relatively few acquaintances. (The logic of sampling error is relevant here: Individuals who employ smaller samples to arrive at a subjective perception of majority opinion are more likely to perceive a majority opinion that varies from the objective majority within the entire population.) In other words, people are most likely to be influenced by a rebel espousing unpopular beliefs if they are acquainted with the rebel but are not acquainted with very many other people. This insight helps to explain why ideologues (with or without disciples) are more successful in spreading their initially unpopular beliefs within populations characterized by lower levels of extraversion. It is within those populations that well-connected ideologues are especially likely to find themselves in the position of being one of very few individuals within their acquaintances' social networks, and therefore able to exert a disproportionately large influence on the beliefs of those relatively lonely acquaintances. However, we don't want to overstate the effect of extraversion. For the realistic range of values in our model, it is interesting that our model predicts that differences in extraversion affect network structure and consolidation (and therefore homogeneity), but have a much weaker effect on the spread of a new idea.

5.5 General Discussion

Results of our computer simulations revealed that cross-cultural differences in individuals' dispositions may have long-term consequences for cultural stability and change. We focused on two empirically-documented cultural differences—differences in mean levels of conformity, and in mean levels of extraversion—and we investigated their implications for two population-level outcomes: (a) The speed with which existing opinion majorities consolidate into even bigger majorities, and (b) the

extent to which initially unpopular beliefs successfully spread within a population. We found that higher mean levels of conformity facilitated the consolidation of majority opinion and also—perhaps counterintuitively—facilitated the spread of unpopular beliefs held initially by a well-connected ideologue (either alone or accompanied by a small number of disciples). Cultural differences in extraversion also had effects on these outcomes: Higher mean levels of extraversion facilitated the consolidation of majority opinion, but inhibited the spread of initially unpopular beliefs.

Although superficially very different, the two population-level consequences of conformity both reflect the same underlying process. One way to think about it is this: An individual's dispositional tendency to conform is equivalent to that individual's likelihood of changing beliefs—of abandoning one belief in favor of another one (as long as it is held by the majority of that individual's acquaintances). Therefore, populations characterized by higher mean levels of conformity are also characterized by relatively greater susceptibility to change. This greater susceptibility for change manifests in the greater likelihood that a small majority will consolidate into a super-majority, and also in the greater likelihood that an initially unpopular opinion will spread. Metaphorically, the mean level of conformity within a population functions like a lubricant: Under conditions in which there exists some potential for cultural change, that potential is facilitated by higher levels of conformity.

The effects of extraversion are somewhat subtler. An individual's dispositional tendency toward extraversion has consequences for the acquisition of acquaintances; this has further consequences for the number of people who are subject to that individual's influence and for the number of influence sources that individuals are exposed to. It is the latter effect that appears to account for extraversion's positive effect on consolidation of majority beliefs and for its negative

effect on diffusion of unpopular beliefs. Because extraverts are exposed to larger samples of people, their subjective perceptions of the majority belief are more diagnostic of the true population-wide majority belief. Therefore, when highly extraverted individuals abandon one belief in favor of the perceived majority belief, they are very likely adopting the true majority belief. By comparison, when highly introverted individuals abandon one belief in favor of the perceived majority belief, they are at greater risk of abandoning the true majority belief and adopting instead an objectively unpopular belief (which just happens to be locally popular among their relatively small number of acquaintances). Thus, in populations characterized by higher mean levels of extraversion, existing majorities become super-majorities more quickly, and radical new beliefs spread more slowly.

Of course, in addition to conformity processes modeled here, other considerations may also affect the consolidation of existing majorities and the diffusion of innovations. Some beliefs are more obviously accurate than others, and some radical new ideas—especially in the realm of technology—are more immediately useful, and so they spread more rapidly for these reasons instead. There may also be additional top-down pressures (e.g., authoritarian governmental policies) that facilitate the spread of some beliefs and inhibit the spread of others. To the extent that this is so, the conformity processes modeled here will be of relatively reduced importance. Therefore, our results probably apply primarily to subjective opinions and beliefs rather than to matters of verifiable fact, and also apply primarily to opinions and beliefs that are relatively unconstrained by laws or other institutional constraints. That still leaves a wide domain of application: These results apply to any idea, opinion, attitude, or behavioral decision that is subject to peer pressure.

5.5.1 Implications for Real-World Populations

Given the importance of the individualism / collectivism distinction in the description of actual human populations, it is interesting to consider the implications that these simulation results

have for predicting differences between individualistic and collectivistic populations in the speed of cultural change. Prototypically individualistic populations are characterized by relatively low levels of conformity and by relatively high levels of extraversion, whereas prototypically collectivistic populations are characterized by high levels of conformity and by low levels of extraversion (Schaller & Murray, 2011).

Our simulation results showed that both conformity and extraversion positively predict the consolidation of small majorities into larger majorities, but also showed that the effect of conformity is substantially stronger than the effect of extraversion. One implication of this difference in effect sizes is that the consolidation of opinion majorities may occur more readily in individualistic populations than in collectivistic populations. An implied difference between individualistic and collectivistic populations is even more evident when considering the speed with which radical beliefs and other innovations spread throughout a population. Our simulation results showed that radical ideas promoted by a single well-connected ideologue were least likely to spread widely within the population that was most prototypically individualistic and most likely to spread widely—and to eventually be held by the majority of people—within populations that were most prototypically collectivistic.

Considered in full, these results imply that individualistic and collectivistic populations may be disposed toward different patterns of cultural change over time. Previous research on the non-linear dynamics of attitude change has suggested that population-level changes in popular opinion may sometimes be described by the mathematics of cusp catastrophes (Latané & Nowak, 1994; Tesser & Achee, 1994). The results of our simulations suggest that the likelihood of this kind of "catastrophic" change differs for individualistic and collectivistic populations. In individualistic populations (characterized by relatively low levels of conformity and high levels of extraversion),

cultural change is predicted to occur slowly, incrementally. By contrast, in collectivistic populations (characterized by relatively high levels of conformity and low levels of extraversion), majorities may more rapidly coalesce into monolithic super-majorities; but when this existing orthodoxy is punctuated by the spread of heterodox beliefs, this change is predicted to proceed at a pace that more closely fits the subjective perception of a “revolutionary” change.

5.5.2 Empirical Testability of the Hypotheses

The results of computer models are *not* empirical observations, of course; they are scientific hypotheses. They represent a set of analytically rigorous predictions about the effects that cultural differences in extraversion and conformity may have on the rate of change in public opinion and popular beliefs over time. Given that these hypotheses pertain to phenomena that pertain to entire populations and must be documented across potentially long stretches of time, they are not easily tested; but they are testable.

These hypotheses may be tested (if not immediately, then eventually) by conducting comparative longitudinal studies on attitudes assessed by survey instruments such as the World Values Survey, which are administered across multiple populations and across multiple periods of time. It might also be possible to test these hypotheses on smaller, geographically-contained populations—such as those that exist in university dormitories or in sororities and fraternities—which can sometimes serve as proxies for larger populations, and have been used previously in naturalistic studies of dynamic social influence and social contagion more generally (Bourgeois, 2002; Crandall, 1988; Cullum & Harton, 2007). It may even be possible to test these hypotheses in laboratory experiments on small groups. Previous research on the cumulative dynamics of social influence processes have attempted to create miniature proxy “populations” in the form of small groups of individuals interacting over small periods of time, with some success (Baum, Richerson,

Efferson, & Paciotti, 2004; Insko et al., 1980; Latané & Bourgeois, 1996; Mesoudi & Whiten, 2008). Similar methods might potentially be used to experimentally (rather than computationally) simulate the variables that define our models, and to test whether conceptually analogous group-level outcomes emerge.

5.5.3 Novel Features of Modeling Methods Employed Here

Of our primary results, only one (the effects of conformity on consolidation of majority opinion) is an intuitively straightforward consequence of individual-level social influence processes. The other results are less intuitive—in part because they emerge from the interplay between the complex geometry of social networks and from the dynamic manner in which interpersonal influence processes unfold over time within those networks. These results highlight the value of rigorous computer models as a means for discovering non-obvious hypotheses about the population-level consequences of individual-level behavioral decisions (Kameda et al., 2003; Kenrick et al., 2003; Latané, 1996; Mason et al., 2007; Nowak, 2004; Pfau et al., 2013; Vallacher et al., 2002).

Computational methods are especially—and perhaps indispensably—useful as a means of identifying the long-term cumulative consequences of interpersonal influence processes. In addition to relevant work within the psychological sciences (Latané, 1996; Nowak et al., 1990), scholars from a wide-range of other scholarly backgrounds (including physics, economics, sociology, and anthropology) have attempted to model social influence processes in order to address a wide-range of topics, including the rise of political extremism (Weisbuch, Deffuant, & Amblard, 2005), changes in consumer preferences (Buenstorf & Cordes, 2008), and cultural evolution more generally (Boyd & Richerson, 1985; Henrich, 2004a; Pfau et al., 2013). Our work here contributes in several novel ways to this scholarly tradition.

The central contribution follows from the fact that our models were designed to simulate individual differences in basic dispositional traits toward conformity and extraversion. Although there may be some general tendency for people to conform to opinion majorities, there is individual-level variability around this central tendency; by simulating this variability, one can model social influence outcomes more realistically. The same principle applies to extraversion. By simulating individual differences in extraversion, we were able to consequently simulate the emergence of social networks with geometric properties mimicking those of actual social networks, thus creating a realistic social ecology within which to examine the cumulative consequences of social influence outcomes. It is worth noting that extraversion has received very little attention in the psychological study of social influence—perhaps because effects of extraversion are not readily apparent on the short-term individual-level influence outcomes that are typically the object of psychological inquiry. But, as our results suggest, extraversion may have effects on the long-term population-level consequences of interpersonal influence.

By simulating individual differences in conformity and extraversion, we were also able to add another novel feature to our models: We simulated cultural differences in conformity and extraversion. This is important because, just as individuals vary around central tendencies toward conformity and extraversion, populations vary in terms of the central tendencies themselves (Bond & Smith, 1996; McCrae et al., 2005). We simulated these cultural differences in a way that mimicked the magnitudes of actual cultural differences documented in the empirical literature. This has useful implications. The predictive utility of a model depends on the extent to which it realistically simulates the variables that are included. By using empirical results to inform our simulations of cultural differences, we can be more assured that the results of our simulations may be sensibly applied to predict outcomes in real human populations.

5.5.4 Lacunae, Limitations, and Directions for Future Research

All models in the behavioral sciences—whether computational or not—represent intentional simplifications of reality. By necessity, these models must omit many of the countless variables that potentially influence individuals' thoughts, feelings, and behavioral decisions. This is not necessarily a limitation (Nowak, 2004). Still, it may be useful to draw attention to some of the specific ways in which our models—like other models of this sort—represent a simplified version of reality, and to consider the implications.

Consider Phase 1 of our models—the phase during which individuals acquired acquaintances and did so in a way that was computationally constrained by geographical proximity and extraversion. While both proximity and extraversion both do have important influences on the formation of social relationships (Asendorpf & Wilpers, 1998; Festinger et al., 1950; Harton & Bullock, 2007; Latané et al., 1995; Paulhus & Trapnell, 1998; Selfhout et al., 2010), other variables matter too. For instance, people are more likely to form relationships with others who have beliefs that are similar to their own (Byrne, 1971)—a tendency that varies in strength across individuals and across populations (Heine & Buchtel, 2009; Schug, Yuki, Horikawa, & Takemura, 2009). The omission of this variable (and of the many variables that can also affect individuals' idiosyncratic decisions regarding who to befriend) did not undermine the objectives of Phase 1—as revealed by results showing that emergent social network structures realistically mimicked the geometric properties of actual social network structures, and also mimicked actual cross-cultural differences in these structural properties. Still, it may be worthwhile in future research to explicitly model both within- and between-population variability in this “similarity-attraction” effect, so as to explore the possible consequences that it too might have on the cumulative consequences of interpersonal influence.

Phase 2 of our models also omitted additional variables that have implications for social influence. In operationalizing the manner in which individuals assess majority opinion, we assumed that all acquaintances' beliefs are treated equally. This is not always the case (in reality, individuals may accord greater weight to the opinions of their parents and siblings than to the opinions of their co-workers or Pilates instructors). More generally, the pool of opinions that really matter may be smaller than the full set of acquaintances that people have. Even if this is the case, however, it has negligible implications for the primary population-level outcomes we observed. The effects of individual differences in conformity are independent of the number of other people whose opinions subjectively matter; and so the effects of cultural differences in conformity will be obtained regardless. And as long as there is some non-zero relation between an individual's level of extraversion and the number of other people that the individual may potentially influence (and be influenced by), then the effects of cultural differences in extraversion will occur as well.

Our simulation of social influence processes also assumed that individuals actually obtain veridical information about others' beliefs. In the real world, this is not always the case. People are sometimes reluctant to express their true beliefs—perhaps especially if they perceive that their beliefs are counter-normative. Indeed, for a variety of reasons bearing on the strategic psychology of social discourse, some beliefs are more likely than others to be the subject of conversations and other forms of interpersonal communication, and these differences in 'communicability' have implications for long-term stability and change in the popularity of these beliefs (Conway & Schaller, 2007; Schaller, Conway Iii, & Tanchuk, 2002). The effects obtained from our simulations pertain primarily to attitudes and beliefs that are communicable in some meaningful way. To the extent that beliefs are less communicable, these effects would be expected to be less apparent.

For the subset of simulations that focused on the diffusion of an initially unpopular belief, we computationally ensured that the primary proponent of that belief was not only ideologically committed, but also highly extraverted. Had we not done so, the baseline likelihood of diffusion would have been substantially reduced, and the effects of both conformity and extraversion would have been reduced accordingly. When interpreting these effects on the spread of a radical new belief, it is important to keep in mind the fact that these effects are specific to conditions in which that radical new belief has some minimally realistic chances of spreading at all.

Across all simulations, we simulated a process in which individuals' are inclined (to varying degrees) to adopt whatever belief is held by a simple majority of their acquaintances. While this is indeed a common decision-rule guiding conformity (Hastie & Kameda, 2005), it is by no means the only such decision-rule. Under different circumstances interpersonal influence may be contingent upon different thresholds of evidence, which may have additional consequences for long-term population-level outcomes (Boyd & Richerson, 1985; MacCoun, 2012). For instance, a more stringent standard of evidence (e.g., a 2/3 majority) would inhibit the speed with initially popular beliefs consolidated and initially unpopular beliefs diffused, and the observed effects of both conformity and extraversion would be somewhat reduced as well. In addition, in our simulations, we conservatively modeled an individuals' likelihood of conformity to be at or below the perceived size of the majority. In many circumstances, the likelihood of conforming exceeds the perceived size of the majority itself—a phenomenon that has been labeled “conformist transmission” (Efferson et al., 2008; McElreath et al., 2005; Morgan et al., 2012). To the extent that the population-level outcomes of dynamic social influence processes are governed by the principles of conformist transmission, it would likely amplify the effects of we observed, for both conformity and extraversion.

Note too that our models were designed to simulate one specific form of social influence: Conformity. While conformity is certainly an important form of social influence (and is the form of influence that is typically simulated in models of consolidation, diffusion, changes in public opinion, and cultural evolution more generally), it is not the only form of social influence. In fact, for psychological reasons that are distinct from those underlying conformity, individuals are sometimes not only *not* motivated to conform, but may actually be motivated to *not* conform to perceived norms (Berger & Heath, 2007, 2008). More broadly, individuals' opinions, attitudes, and beliefs also change in response to persuasive messages—many of which are crafted with considerable cunning to take advantage of psychological processes that are independent of those that affect conformity, but which may still affect attitude change (Albarracín & Vargas, 2010). To the extent that these additional psychological processes also influence the consolidation of belief majorities and the diffusion of new beliefs, they represent phenomena that are conceptually independent of those examined by our models, and would need to be simulated separately in future models.

Finally, while our models are the first to rigorously examine the effects of dispositional variability (both within and between populations) on dynamic social influence outcomes, we focused on just two of the many dispositional differences that may have implications for social influence processes. Other individual difference variables may matter too. For instance, within the psychological literature on persuasion processes, there is evidence that the influential impact of persuasive communications may be moderated by individual differences in needs for cognition and for cognitive closure (Cacioppo, Petty, Feinstein, & Jarvis, 1996; Kruglanski, Webster, & Klem, 1993). Not only do individuals vary in the extent to which they chronically experience these epistemic needs, there are cultural differences too (Chiu, Morris, Hong, & Menon, 2000). What implications might these individual and cultural differences have on the cumulative population-level

consequences of interpersonal persuasion? We do not know. In order to sensibly speculate, it will be necessary to develop new models that, while conceptually distinct from our models (which focus on conformity rather than persuasion processes), incorporate analogous methodological innovations. For example: It may be possible to realistically simulate individual (and cultural) differences in need for cognitive closure, and also simulate the effects that these differences have on persuasion processes, and in doing so, computationally assess their long-term population-level consequences.

5.5.5 Broader Applications of These Modeling Methods

As the preceding paragraphs illustrate, the modeling methods that we have used are flexible, and can be amended to address additional interesting questions about effects of cultural differences on the population-level consequences of interpersonal influence. Our modeling methods may have a broader set of useful applications as well.

For example, the methods we used to simulate the emergence of realistic social network structures (in Phase 1 of our simulations) might be profitably amended to model the effects that other variables have on emergent social network structures, and to examine the consequences. Populations are typically comprised of people defined by different demographic categories (gender, ethnicity, language, etc.); these differences affect the formation of relationships that, in turn, affect a wide range of outcomes of considerable psychological and societal importance—including prejudice and the acculturation of immigrants (Laar, Levin, Sinclair, & Sidanius, 2005). The processes can be formalized with the modeling methods that we employed, allowing for rigorous exploration of emergent population-level consequences of demographically constrained patterns of friendship formation (Pfau et al., 2013).

These modeling methods might also have useful applications in the study of group decision-making. Although we have applied these methods to research questions bearing on large

populations, the methods can be easily amended to address research questions pertaining to smaller groups (Hastie & Kameda, 2005; Kerr & Tindale, 2004). For example, recent research shows that the effect of group size on the quality of group decisions depends on the extent to which group members make independent intellectual contributions to these decisions (Kao & Couzin, 2014). The independence of individuals' contributions is itself likely to depend, in part, on the group's social network structure—which, as we have shown, is influenced by the dispositional traits of group members. With minor amendments, our modeling methods might profitably be used as a means of identifying hypotheses about the effects that individual differences, and cultural differences, may have on group decision-making.

These methods may also have useful applications within the multi-disciplinary study of cultural evolution. Although there are many sophisticated models of cultural evolution (Boyd & Richerson, 1985; Henrich, 2004a), it is rare for these models to explicitly simulate the geometric properties that define the social network structures of real human populations. For example, Chapter 3 reveals relationships between individual-level sociality and emergent cultural complexity (Muthukrishna, Shulman, Vasilescu, & Henrich, 2013); however, these results were based on models that—like most cultural evolutionary models—made simplifying assumptions about social network structure governing the interpersonal transmission of cultural information. By incorporating the methods employed in Phase 1 of our models, it may be possible to ask, and answer, questions about the realistic effects of social network structure on cultural transmission and cultural evolution.

5.5.6 Envoi

There is a substantial body of computational modeling research identifying the population-level consequences of interpersonal influence outcomes as they accumulate dynamically across time (Axelrod, 1997; Mason et al., 2007; Nowak et al., 1990; Valente, 1995); but no prior research within

this tradition had addressed questions about cultural differences on these influence outcomes. There is another substantial body of empirical research documenting effects of culture on social influence phenomena (Bond & Smith, 1996; Kim & Markus, 1999; Zou et al., 2009); but that research has focused almost exclusively on short-term individual-level outcomes. Our work represents a conceptual bridge between these two scholarly literatures. In doing so, it makes novel conceptual contributions to the psychological study of social influence and its cumulative consequences, and also to the study of cultural differences. Also, by showing how individuals' actions create specific kinds of ecological circumstances (e.g., social network structures governing patterns of interpersonal interaction), and showing how those ecological circumstances consequently affect individual- and population-level outcomes, this work also contributes to an emerging literature on socioecological psychology (Oishi, 2014). More broadly, it contributes both methodologically and conceptually to multi-disciplinary inquiry into the dynamic processes through which ideas spread, norms change, and populations evolve.

Chapter 6: Conclusion

During my time in graduate school, a “crisis” emerged in my field. One of the watershed moments included Daryl Bem publishing statistical evidence for Extrasensory Perception (ESP; Bem, 2011) using research practices and statistical methods that many felt were routine in psychology and certainly not objectionable enough to warrant a rejection from the field’s flagship journal. A year later, Daniel Kahneman wrote an email to colleagues calling for more attempts to replicate priming studies and warning of a “train wreck looming”. The email was republished in *Nature News* (Yong, 2012). Such concerns have been simmering for a while¹⁶, but with events such as these, the pot overflowed. The field responded with clearer identification (and disapproval) of researcher degrees of freedom (“p-hacking”; Simmons, Nelson, & Simonsohn, 2011), new methods to detect such violations (Simonsohn, 2013; Simonsohn, Nelson, & Simmons, 2014), routine reporting of effect sizes and confidence intervals (Cumming, 2013), large-scale replication attempts (Yong, 2013), pre-registration of studies similar to practices in pharmaceutical research (Nosek & Lakens, 2014), and badges for good practices (Eich, 2014). The discussion continues as I write this dissertation and at least one major journal (*Psychological Science*) has changed its rules with respect to statistical practices and reporting research practices.

¹⁶ Examples include criticisms of null hypothesis testing and discussions of Bayesian vs frequentist approaches, including Jacob Cohen’s classic “The earth is round, $p < .05$ ” (Cohen, 1994) and the edited volume “What if there were no significance test” (Harlow, Mulaik, & Steiger, 1997); discussions about the practice of HARKing – Hypothesizing After the Results are Known (Kerr, 1998) and what Charles Peirce called “abductive” reasoning as an alternative to the hypothetico-deductive approach to science (e.g. Rozeboom, 1997); Walter Mischel describing what he called “The Toothbrush Problem” (Mischel, 2009) where researchers avoiding using other researcher’s theories as they would avoid using other people’s toothbrushes and discussions on common methods and how to make psychology a cumulative science (e.g. *Psychological Methods* released a special issue titled “Multi-Study Methods for Building a Cumulative Psychological Science” in June, 2009).

The approach taken in this dissertation offers a different perspective to the common consensus. There are statistical issues in the field, mainly centered around researcher degrees of freedom and the abuse and misuse of statistical methods, which some of the suggestions listed above may help resolve, but at its core, our field's problem is one of theory.

6.1 A Theory of Human Behavior

The advantage of a good theory is that it not only makes predictions about what to expect, but also exclusions about what not to expect. As Karl Popper (1962) puts it, “the more a theory forbids, the better it is.” (p. 36). With our scientific intuitions tuned by theory rather than life experience, we're better able to identify when something seems “off”. When neutrinos appeared to be travelling faster than the speed of light (Agafonova et al., 2012), physicists knew something was wrong, because it violated the Theory of Special Relativity. If vinegar (acetic acid) and baking soda (sodium bicarbonate) combined in your child's model volcano doesn't produce carbon dioxide and hot ice (sodium acetate) solution, chemists know something is wrong, because it violates the Periodic Table and Collision Theory. If fossil rabbits were found in the Precambrian era, biologists would know something was wrong, because it violates the Theory of Evolution. But if humans seem to prefer less choice to more (Schwartz & Kliban, 2004), does this violate our Theory of Human Behavior? What if humans prefer more choice to less or show no preference at all (Scheibehenne, Greifeneder, & Todd, 2010)? If humans appear to walk slower when they're reminded of old people (Bargh, Chen, & Burrows, 1996) does this violate our Theory of Human Behavior? What if they walk faster or ambulate unperturbed by memories (Doyen, Klein, Pichon, & Cleeremans, 2012)? Without an overarching scientific Theory of Human Behavior from which to draw hypotheses and tune our intuitions, it can be difficult to distinguish results that are unusual and interesting from results that are unusual and likely wrong.

There are explanations for the two examples I offered – the choice overload and elderly behavioral primes – and these explanations do come from a Theory of Human Behavior. But that Theory emerges from each researcher’s own life experience and perhaps past experimental data. We might call these explanations theories or hypotheses, but as theories they lack generality in predictive power and as hypotheses they flow from each researcher’s culturally specific intuitions (Henrich, Heine, & Norenzayan, 2010) rather than an overarching theory.

Mini-theories and hypotheses based on intuitions or past data are not necessarily a problem. In an applied context, such as pharmaceutical trials, testing a drug and showing its efficacy works regardless of the drug’s origins¹⁷. In the applied science of declaring drugs effective for human use, the pharmaceutical sciences have established useful best practices such as multiple studies and pre-registration of methods, sample sizes, and analyses. These practices help establish the presence and size of an effect and prevent changing hypotheses after seeing results (Kerr, 1998). In a basic science context, in principle these hypotheses or mini-theories could coalesce into a larger overarching theory, but in practice avoidance of others mini-theories (Mischel, 2009) and a lack of common methods can slow down or prevent the cumulative process. Moreover mini-theories, especially if they are not formally specified, lend themselves to confirmation rather than falsification and as Popper (1962) points out, “It is easy to obtain confirmations... for nearly every theory—if we look for confirmations” (p.36)¹⁸. One key advantage to a Theory of Human Behavior in the quest for a

¹⁷ Examples of such origins include traditional knowledge (St John’s wort), past side effects (Viagra), similar chemical compounds (Captopril was the first ACE inhibitor heart medication developed, but others such as perindopril and ramipril soon followed).

¹⁸ Karl Popper (1962) (p.36) lists 7 criteria for a scientific theory, quoted below.

1. It is easy to obtain confirmations, or verifications, for nearly every theory—if we look for confirmations.
2. Confirmations should count only if they are the result of *risky predictions*; that is to say, if, unenlightened by the theory in question, we should have expected an event which was incompatible with the theory—an event which would have refuted the theory.

cumulative science is that it allows us to interpret past findings in the way that the Periodic Table and Collision Theory allow you interpret pre-Mendeleev chemical experiments.

Psychology has decades of data gathered using clever methods and manipulations, but with that data now under suspicion¹⁹, a Theory of Human Behavior would be a useful way to parse which results are most suspicious and thereby move forward as a cumulative science. But the mere fact that such an overarching theory would be useful, does not imply that one exists or that any will suffice. There are many candidate theories. A popular Theory of Human Behavior in economics is that of economic man or Homo Economicus, a theory borne out of 19th century philosophy (Persky, 1995). Homo Economicus conforms to the requirements of a scientific theory and is going through a process of improvement after some of its predictions have been challenged (Gintis, 2000; Henrich et al., 2001; Kahneman & Tversky, 1979; Thaler, 2000). I suspect that the most predictive Theory of Human Behavior will flow from the Theory of Evolution, connect with the growing body of knowledge in neuroscience and genetics, and explain both cross-species differences and cross-

-
3. Every 'good' scientific theory is a prohibition: it forbids certain things to happen. The more a theory forbids, the better it is.
 4. A theory which is not refutable by any conceivable event is nonscientific. Irrefutability is not a virtue of a theory (as people often think) but a vice.
 5. Every genuine *test* of a theory is an attempt to falsify it, or to refute it. Testability is falsifiability; but there are degrees of testability: some theories are more testable, more exposed to refutation, than others; they take, as it were, greater risks.
 6. Confirming evidence should not count *except when it is the result of a genuine test of the theory*; and this means that it can be presented as a serious but unsuccessful attempt to falsify the theory. ...
 7. Some genuinely testable theories, when found to be false, are still upheld by their admirers—for example by introducing *ad hoc* some auxiliary assumption, or by re-interpreting the theory *ad hoc* in such a way that it escapes refutation. Such a procedure is always possible, but it rescues the theory from refutation only at the price of destroying, or at least lowering, its scientific status.

One can sum up all this by saying that *the criterion of the scientific status of a theory is its falsifiability, or refutability, or testability.*

¹⁹ The latest draft of the Open Science Framework's Many Labs replication effort found that only 3 of the 10 effects tested replicated: <https://osf.io/s59bg/>

cultural variation. It is such a theory that is extended (Chapter 2), tested (Chapters 3 and 4) and challenged with suggested improvements (Chapter 5) in this dissertation.

Humans are an evolved species and Dobzhansky's (1973) famous phrase applies as much to our psychology as to our biology – “Nothing makes sense except in the light of evolution”. Like all other species on the planet, all aspects of our behavior must flow from the evolutionary processes that led to our present state. My dissertation builds on and tests one candidate evolutionary Theory of Human Behavior – Dual Inheritance Theory (or Gene-Culture Coevolution) – the idea that the same evolutionary processes that led to every species on the planet led humans down a unique pathway. Selective forces, some of which are described in Chapter 2, led to the development of a suite of psychological abilities and tendencies that allowed our species to learn from each other with high fidelity. This high fidelity learning led to a second-line of inheritance – culture. Genes adapted to this new selection environment, which now included culture and those genes in turn enabled new cultural information in a co-evolutionary process.

From a Dual Inheritance Theory perspective, the research enterprise involves understanding how evolution shaped our brains and bodies in ways that allowed us to acquire culture (Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981), identifying what psychology is required for culture and how culture emerges from that psychology (e.g. Henrich & McElreath, 2003; Schaller & Crandall, 2003), how culture mutually shapes our brains and bodies (e.g. Laland et al., 2010), and how culture itself evolves and leads to cross-cultural differences and societal-level phenomena (e.g. Heine & Norenzayan, 2006; Henrich & Boyd, 2008; Henrich, Boyd, & Richerson, 2008; Norenzayan & Heine, 2005; Schaller & Murray, 2008). Dual Inheritance Theory and Cultural Evolution offer several predictions and exclusions for the psychology of our species. The chapters of this dissertation contribute to this enterprise, both in theory and in tests of theory.

6.1.1 Building Theory

The best scientific theories make general predictions and exclusions about what to expect and not expect and are more parsimonious than alternatives. These theories can be expressed in many ways. Natural Selection as a Theory of Evolution was first expressed as a verbal argument (Darwin, 1859). Evolutionary theory has come a long way in the century and a half since Darwin published his classic²⁰. For example, we now know about genetics and have a much better understanding of the many processes through which species diversify and evolve. Since the Modern Synthesis, evolutionary biology has expressed its theories using mathematical and computational models. There are good reasons for why this is a useful tool for theory building.

Researchers use formal mathematical and computational models in all kinds of ways – high fidelity simulations of reality (MSC Software, 2004), precise quantitative predictions of systems like the stock market (Chan, 2009), probabilistic models for tasks like facial recognition (Liu & Wechsler, 2002) and so on. Unlike in these cases, biologists, anthropologists, and psychologists often use formal models as aids to thinking through the logic of an argument in order to make testable *qualitative* predictions about phenomena (e.g. Aoki & Feldman, 2014; Boyd & Richerson, 1985; Hastie & Kameda, 2005; Kendal, Giraldeau, & Laland, 2009; MacCoun, 2012; Nowak et al., 1990; Tanford & Penrod, 1983, 1984). By formally defining assumptions, logic, and predictions, anyone can challenge the theory by either testing the predictions or by modifying the assumptions or logic. By deciding on the minimal set of assumptions required to explain a phenomena and formally expressing these assumptions, the logic that follows, and the predictions and then modifying assumptions, logic, and predictions in the face of empirical evidence, we can start to build a

²⁰ Although perhaps because our tendency to rely on prestigious figures (Chudek, Heller, Birch, & Henrich, 2012; Henrich & Gil-White, 2001), researchers still tend to use Darwin's words to bolster their case.

cumulative science. And expressing these in the language of mathematics has some advantages in the quest for a cumulative science.

Why not rely on verbal arguments? Well, for simple if-then causal relationships, you can get away with words. For example, when the sun is out people eat more ice cream. And perhaps this is mediated by temperature. But our minds are limited in memory and processing (Gigerenzer & Selten, 2002; Kahneman, 2011; Miller, 1956) so arguing with words or just thinking through a theory – effectively simulating in our minds – gets fuzzy fast. In their now classic 1985 book, Robert Boyd and Peter J Richerson express the choices of how to express a theory as, “the real choice is between an intuitive, perhaps covert, general theory and an explicit, often mathematical one... Many aspects of a scientist’s mental model are likely to be vague and never expressed” (p.27). Thankfully, we have a couple of cultural technologies that allow us to overcome these mental limitations: analytic models – solvable systems of equations – and when it gets more complicated, computational models. By building formal testable theories to explain the world, we can test competing theories, and bring the theories of the biological, psychological, anthropological, and human evolutionary sciences into a broader scientific framework.

Since *Homo habilis* (or perhaps *Australopithecus*; Harmand et al., 2015) first banged two rocks together to make a chopping tool, specialized tools have allowed us to overcome the limitations of our bodies. Hammers let you hit harder; trains let you travel further. In modern societies, many tools are instrumental in overcoming the limitations of our mental faculties. The simple pen and paper let you remember more, computers let you calculate faster. Most hypotheses in the psychological sciences are generated without the need for any such specialized tools, because the typical objects of inquiry (unidirectional causal relations operating at a single level of analysis) are amenable to informal logical deduction. Although, as I have argued, there are advantages to a Theory of Human

Behavior, especially in a basic science context. But even without a broader Theory of Human Behavior, when addressing questions about phenomena defined by more complex causal relations that play out dynamically over time and produce emergent consequences that must be measured at a different level of analysis entirely, specialized tools *are* needed. Mathematical and computational models can be thought of as aids to thinking, allowing us to work through the logic and assumptions of systems more complex than our minds can fully represent. Models need to be abstract enough to be more tractable than reality, but realistic enough to sufficiently capture the problem and inform our understanding of it. They allow us to make formal, precise predictions that go beyond the imprecision of words. But models are only useful insofar as their logic and assumptions are informed by empirical research, and in turn they make predictions that can be tested empirically.

6.1.2 Testing Theory

General unifying theoretical frameworks like Dual Inheritance Theory are not monolithic or complete. The details are worked out by generating “sub-theories” and testing them. These formal theories make specific predictions; different theories make competing predictions. To distinguish between competing theories, we must turn to empirical data.

There are many ways we can test theoretical predictions. Each method has its pros and cons. Laboratory experiments, used in both Chapters 3 and 4, give you control at the expense of true ecological validity. In field experiments, you lose some of that control, but gain ecological validity. Finally, there are existing data-based methods, such as those used in Chapter 2 and suggested in Chapter 5. However, without randomized controlled trials, causality is more difficult to infer and at best we can say that they do not falsify the theory. Together these methods are best deployed not to

confirm a theory, but in a good Popperian manner test competing theories and hopefully falsify one²¹.

6.1.3 Present Research Enterprise

The approach I have discussed thus far is the approach used throughout this dissertation. In Chapter 2, I challenged the most popular explanation for the evolution of large brains – The Social Brain Hypothesis (Dunbar, 1998). I presented a formal model and argue that the Social Brain Hypothesis is actually part of a larger more general process, which I call the Cultural Brain Hypothesis. The Cultural Brain Hypothesis model makes several predictions, which I test using existing data. These results reveal that the Cultural Brain Hypothesis can explain the same evidence as the Social Brain Hypothesis, but also additional evidence that has yet to be explained by a single theory. I therefore argue that the Cultural Brain Hypothesis is a more general and parsimonious explanation to alternative theories. Further, under some conditions the same mechanisms underlying the Cultural Brain Hypothesis also lead to a separate evolutionary pathway where culture accumulates and exerts further selection pressures on large brains, which in turn allow for more culture. The conditions that lead to this autocatalytic coevolution are the predictions of the Cumulative Cultural Brain Hypothesis. The predictions are consistent with other more specific models and with empirical data, including the data presented in Chapters 3 and 4.

Different formal theories may make the same predictions, but have different assumptions and logic underlying those predictions. Thus they may both be consistent with existing correlational

²¹ Of course, one falsification is not enough to overturn a theory. And theories can and should be modified until they can stand no more against the weight of contrary evidence. Even so, scientists are humans and scientific revolutions rarely proceed in this ideal manner. Instead scientific revolutions may go through Kuhn's 5 phases and perhaps require a generational shift. As Max Planck is reported to have said: "A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it" (Kuhn, 1962).

data and only separable by an experiment. This is the case for the relationship between sociality and cultural complexity, tested in Chapter 3. Take the case of cultural loss following population loss. Cultural drift style models predict that this loss is caused by a process analogous to genetic drift, whereby there are fewer people to remember the culture. On the other hand, “treadmill” style models predict that this loss is due to a loss in sociality – the size and interconnectedness of populations – which is required to compensate for imperfect transmission fidelity. Our results were consistent with a treadmill style model.

Theories also shape the way in which experiments are designed and can help explain why they fail to show results. Muthukrishna et al. (2013), reported in Chapter 3, was not the first experiment to test these competing theories. Caldwell and Millen (2010) previously showed no effect of number of models on mean skill level in a paper airplane task. The impetus for the studies reported in Chapter 3 was that their study failed to capture some necessary aspects of the theory. Specifically, the theoretical models Caldwell and Millen (2010) tested predicted that if some skill or other cultural trait is sufficiently easy to learn or cognitively transparent, then increasing the number of models available to learners will have little impact (Henrich, 2004b). That is, making a simple paper airplane is too easy to learn and their experiment was therefore not a test of the competing theories. When we used tasks that were difficult for an individual to learn in one lab generation (image editing and tying a system of rock climbing knots), the results supported a treadmill model over a drift model.

In Section 6.1, I used choice overload as an example of a finding that does not seem to emerge from a well-specified Theory of Human Behavior. I deliberately chose this example, because of its relevance to the experiments reported in Chapter 4. In two experiments, I tested several social learning predictions that emerge from Dual Inheritance Theory and Cultural Evolutionary models.

These included predictions about the relationship between number of options and social learning. Without a well specified theory, the question asked is “whether more choice is better” where “better” can be operationalized in a variety of perfectly acceptable ways (Schwartz & Kliban, 2004). These are fine applied questions of particular relevance to marketers. But from a basic science perspective, in the enterprise of building an edifice of knowledge about the psychology of our species, starting from the bottom up in trying to understand these phenomena can be difficult, because the question might be unanswerable. Is walking or running better? Even if we specify what better means (the easy part) – let’s say, for travelling from home to work – the question may still be unanswerable. It depends on so many other factors. How might we narrow down the large or infinite set of possible predictive factors to build a theory – individual differences like fitness, external factors, like how quickly one needs to get to work, alternative options like driving? This is the essence of abductive reasoning. Even if we were to show that on average people prefer running to walking to work, it can be difficult to determine what factors matter when results don’t replicate. Indeed, in the case of choice overload, a meta-analysis suggested no effect of choice overall, but admitted that there may be unspecified moderators driving the effect, if it exists (Scheibehenne et al., 2010). In contrast, a Theory of Human Behavior, especially one derived from larger established theories (such as the Theory of Evolution) not only shapes how to test predictions, but also the question that is asked. From a Dual Inheritance Theory perspective, Scheibehenne et al. (2010) are right in that there are many factors beyond number of choices that determine if “more choice is better”. A more appropriate starting point is that humans have faced more and less choice over their evolution. They have a suite of psychological tools to deal with that choice depending on the importance (payoffs) and immediacy (time constraints) of the decision, and what information they have available (individual learning, social information, environmental information, choices

themselves, and so on). A more appropriate question is how humans behave with more and less choice under different conditions. Which conditions are important and which are not, can be derived via formal evolutionary theory. Nakahashi et al. (2012) model how people react to different sized majorities depending on the number of choices available. Their theory and our empirical results suggest that people are more biased towards majorities as the number of options increases. Our empirical results also suggest that they are more likely to defer to use social information to make that decision. Nakahashi, Wakano, and Henrich's (2012) model is a first attempt to formally model decision-making with multiple choices and there are of course many other important theoretical predictions to be made in this area. A formal Theory of Human Behavior shapes the relevant question and narrows down the search space of important factors.

Chapter 4 is also a good example of experiments finding results that go beyond theory. We started with theory and designed an experiment that captured different aspects of the problem, such as the effect of number of options in Experiment 1. Our results reveal a pattern of conformist biased social learning that closely matches theoretical predictions. But our results also reveal the effect of asocial priors based on individual learning, the effect of number of options on the rate of social learning overall, and the effect of individual differences, and in the case of transmission fidelity (Experiment 2); results contrary or not captured by the theory. All of these point to ways in which the theory needs to be reconciled.

Finally, Chapter 5 is also an example of empirical data challenging theoretical models, but in this case based on assumptions rather than predictions. One simplifying assumption made in many Dual Inheritance Theory models, including the model reported in Chapter 2 is a uniform social structure – i.e. all individuals have roughly the same number of connections. While it may be that network structure is irrelevant to the predictions of these theories, the real world shows some

consistent network structures across societies at different scales. These structures may affect human behavior and population-level outcomes and have been shown to affect the efficiency of information transmission (Pasquaretta et al., 2014). One barrier to incorporating these structures into formal theory is a lack of theory to explain how these structures emerge from individual decision-making. In Chapter 5, we introduce a potential theoretical explanation for these network structures and show ways in which they affect population-level outcomes – the consolidation of majorities and the spread of innovations. The effect of these network structures on these outcomes indicates that they may also affect other outcomes of relevance to psychologists, anthropologists, and biologists. Chapter 5 is a first step in this direction.

The Theory of Human Behavior that this dissertation relies on is based on the idea that our basic psychology flows from our capacity for culture. Many aspects of this theory are modeled in Chapter 2. There remain several central questions underlying the evolution and psychology of our species. Some of these have answers in theory, others supported by data, and still others that are poorly understood. In the next section, I lay these out and in doing so, situate the present dissertation.

6.2 Central Questions and Answers

Thus far, I have described one potential answer to why humans are so different to other animals: Humans have a second line of inheritance – cultural inheritance – that has accumulated over generations and shaped our psychology and physiology. This answer opens up further central questions and answers.

6.2.1 Why Now?

Life first appeared on this planet approximately 3.7 billion years ago (Ohtomo, Kakegawa, Ishida, Nagase, & Rosing, 2014). Humans and chimpanzees went their separate ways approximately

4-5 million years ago, the earliest stone tools have been dated at approximately 2.6 million years ago (Semaw et al., 2003) and the particular branch of humans that replaced all others (ours) emerged around 200-300 thousand years ago (Scally & Durbin, 2012). If a second line of inheritance is the key to the human success story, why did it only appear in less than the last 0.1% of the history of life²²? Like with most of these central questions, research is ongoing, but Dual Inheritance Theory predicts that part of the answer lies in temporal and environmental variation (Aoki & Feldman, 2014; Boyd & Richerson, 1985; Nakahashi et al., 2012). If the environment is stable over space and time, slowly evolving genes can adapt to the environment without the need (and cost) of high cognition. If on the other hand, the environment is highly unstable, the information possessed by previous generations or settled groups is less useful and individual learning is needed to adapt to the changed environment. But between these two extremes is a Goldilocks zone where some amount of social learning is beneficial. With a moderately stable/unstable environment, genes are too slow to adapt, and individual learning is more expensive than simply learning from the adapted previous generation (temporal variability) or others already in the environment (migrants facing spatial variation). It is here that the second line of inheritance can evolve.

Testing these theoretical predictions can be difficult, but in recent years ice cores have provided us with high resolution climate data. Martrat et al. (2007) data reveals that at least in the last 420,000 years, climatic variability has increased, consistent with Boyd and Richerson's (1985) model. Exactly the kind of climate variation one would expect for a cultural species to emerge. So the climatic variability was perfect for social learning and culture to be favored, but humans weren't the

²² Speculated sapient dinosaurs aside (Russell & Séguin, 1982).

only species affected by that climatic variability, so the next central question is why did cumulative culture evolve in our species and not others?

6.2.2 Why Us?

The first answer to this question is that our species is not alone in the evolution of social learning. Social learning is widespread in the animal kingdom (Hoppitt & Laland, 2013; Laland, 2008; Whiten & Van Schaik, 2007). Moreover, social learning is positively correlated with brain size (Reader & Laland, 2002) and brain size has been increasing across many taxa (Shultz & Dunbar, 2010a). Chapter 2 presented the Cultural Brain Hypothesis, a theoretical explanation for this encephalization and other associated patterns. But of course, while other species may have had similar selection pressures for larger brains and even culture, humans are unique in possessing *cumulative* culture. The answer to why humans are alone in our domination of the planet via the capacity for culture likely lies in proto-human physiology, psychology, and sociality. The complete and precise physiological pre-requisites for a cultural species have yet to be theoretically or empirically expressed. In the meantime, we may speculate. For example, primates have an advantage over intelligent birds (Emery, 2006) and dolphins (Whitehead & Rendell, 2014) in having hands, which are useful for manipulating objects and eventually making and carrying tools. Thus hands, a necessity for arboreal life, may have been exapted once humans began on the path to a cultural species. Hands may have also been useful (though not necessary) for gestural communication, which eventually led to language (Gentilucci & Corballis, 2006). But hands can't be the only pre-requisite – we share them in common with other primates. Bipedalism is another candidate pre-requisite. Bipedalism may have evolved as early as 4.4 million years ago (White, Suwa, & Asfaw, 1994) and together with hands, may have allowed our ancestors to freely communicate with gestures and to make and carry tools over large distances. These and other physiological pre-requisites may have

been useful or even necessary for making the transition to a fully cultural species. Most of the later changes were psychological and social and it is here that my dissertation makes its contributions.

6.2.3 What Psychology Do We Need For Culture?

What are the psychological and social foundations of culture? The Cumulative Cultural Brain Hypothesis (Chapter 2) captures some of the psychological and social requirements for the leap from a cultural to a cumulative cultural species. Specifically, it makes 4 predictions: (1) high transmission fidelity, (2) low reproductive skew and/or cooperative breeding, (3) smart individual learning ancestors, and (4) an ecology that can be exploited by more knowledge. There are many psychological and social factors underlying each of these predictions.

The Cumulative Cultural Brain Hypothesis is not alone in highlighting the importance of high fidelity cultural transmission (Claidière & Sperber, 2010; Lewis & Laland, 2012). Several experiments in the psychological sciences have demonstrated that humans have a tendency to *overimitate* – copy with high fidelity (Nielsen, Subiaul, Galef, Zentall, & Whiten, 2012; Over & Carpenter, 2013; Whiten, McGuigan, Marshall-Pescini, & Hopper, 2009). High fidelity transmission has at least 3 requirements: an ability, a proclivity, and social infrastructure.

The ability to copy information with high fidelity is likely supported by other cognitive mechanisms like theory of mind. Theory of mind may have evolved for non-cultural reasons, such as for dealing with the complexities of life in larger groups. Chimpanzees, for example, do seem to have some components of theory of mind, although they lack full-blown human theory of mind (for a review, see Call & Tomasello, 2008). Despite missing some components of human theory of mind (which may be irrelevant in a chimp world), chimpanzees have shown some ability to imitate (Whiten et al., 2009). Imitation ability is a necessary, but not sufficient requirement for cumulative cultural transmission. A species also needs a proclivity to do so.

Horner and Whiten (2005) first showed that children have a tendency to imitate even causally irrelevant actions. This result has been replicated under different circumstances (Lyons, Young, & Keil, 2007; McGuigan, Whiten, Flynn, & Horner, 2007) and with adults (Flynn & Smith, 2012; McGuigan, Makinson, & Whiten, 2011). Humans, at least human children, are also often selective in when, who, and what they imitate, and understand the difference between goal-driven or conventional actions (Herrmann, Legare, Harris, & Whitehouse, 2013). The human proclivity for high fidelity transmission is supported by various other characteristics of our species.

Humans are social (Boyd & Richerson, 2009) and prosocial (Bell, Richerson, & McElreath, 2009; Chudek & Henrich, 2011), giving learners exposure to several potential models. For more difficult tasks, these models go as far to slow down their actions or even teach (Fogarty, Strimling, & Laland, 2011; Kline, 2014). And children expect this, inferring that more knowledgeable models are also more prosocial (Brosseau-Liard & Birch, 2010)!

Human life history also supports the transmission process. First via an extended juvenile period in which additional learning may take place (Gurven, Kaplan, & Gutierrez, 2006; Henrich, forthcoming; Joffe, 1997). If adolescence is defined as the period between sexual maturity and reproduction, there is some evidence that this period is extending even further (Mathews, Hamilton, & National Center for Health Statistics, 2009), perhaps through cultural evolutionary processes. Second, the transmission process is supported by a long lifespan and post-menopausal period, where females (and perhaps males) serve as repositories of accessible knowledge – an Information Grandmother Hypothesis. There is some evidence of this in orca, another highly cultured species. Orca grandmothers lead hunting groups, particularly when resources are scarce (Brent et al., 2015) and their presence increases the survival of their sons (Foster et al., 2012). In humans, Henrich and Henrich (2010) have suggested this hypothesis in the Supplemental Materials of their paper, which

showed evidence that Fijian women acquire their adaptive food taboos from their grandmothers. For additional discussion, see Chapter 8 of Henrich (forthcoming).

6.2.4 What Sociality Do We Need For Culture?

Access to multiple models, and perhaps multiple generations, is necessary for the third aspect of an evolutionary system – variation reduction. “One cultural parent makes no culture” (Enquist et al., 2010). The Cumulative Cultural Brain Hypothesis predicts that lower reproductive skew is more conducive to the entering the realm of cumulative cultural evolution, allowing for more genetic variability. One social structure that may have served both low reproductive skew and easier access to multiple models is cooperative breeding. Several researchers have posited the existence of ancient cooperative breeding human societies (Emlen, 1995; Hrdy, 2009; Kaplan, Gurven, Hill, & Hurtado, 2005; Kaplan, Hill, Lancaster, & Hurtado, 2000; Mace & Sear, 2005; Wiessner, 2002) with some evidence of cooperative breeding among modern hunter-gatherers (Hill & Hurtado, 2009). The suggestion (which to my knowledge, has not been formalized) is that a young proto-human primate may have initially learned from mom as many chimpanzees do today (Boesch, 1991; Lind & Lindenfors, 2010; Tagliacola, Reamer, Schapiro, & Hopkins, 2012). Mom may be the primary model simply because her children spend more time with her. Cooperative breeding may have provided a young proto-human access to more moms (and perhaps dads); a gateway to biased social learning, where a young learner could focus on characteristics of the models rather than how much access they had to them. In the model presented in Chapter 2, once social learning evolves, there is a selection pressure for oblique learning to take advantage of other models in the group and learning biases to select from these. The experiments in Chapter 3 also highlight the importance of sociality – the size and interconnectedness of populations – in the evolution and accumulation of

culture. Individuals with access to more models appeared to learn from the best model and then integrate further information from the next two best models.

6.2.5 How Does This Connect With Our Broader Psychology?

The psychological sciences have a wealth of data on the ways in which information is transmitted between individuals; i.e. cultural transmission. These data are often tests of Dual Inheritance Theory and Cultural Evolution predictions and in other cases can inform these theories (Mesoudi, 2009). Some areas of research within social psychology that are of particular relevance to cultural transmission include conditioning (operant conditioning, classical conditioning), social learning (Bandura, 1977), social influence, including norm psychology (Bond & Smith, 1996; Cialdini & Goldstein, 2004; Moscovici, 1980), persuasion, including attitude change (Albarracín & Vargas, 2010; Kumkale & Albarracín, 2004; Petty & Briñol, 2011), and social cognition more generally (Fiske & Taylor, 2013). Other relevant areas include research on the psychology of leadership (Van Vugt & Ahuja, 2011) and group dynamics (Hogg, 2013). Within cognitive psychology, two particular areas of interest are the psychology of language (Traxler & Gernsbacher, 2011), mental models and schemas, and cognitive biases (Kahneman, 2011). The developmental trajectory of these psychologies are also studied within developmental psychology (e.g. mental models; Legare & Clegg, 2015; language; Werker & Hensch, 2015). Cultural psychology (Kitayama & Cohen, 2010) reveals some of the variability in human psychology and the predictors of these differences. Finally, evolutionary psychology has identified potential genetically evolved biases (e.g. detecting cheaters (Cosmides & Tooby, 1992), parental care motivations (Buckels et al., 2015), sex differences in mating choices (Miller, 2011)). In some cases, these offer a competing Theory of Human Behavior (for other examples, see Laland & Brown, 2011). Recent efforts have sought to unify this line of research with

evolutionary theory more broadly, including Dual Inheritance Theory and Cultural Evolution (Barrett, 2014).

There is much work to be done in systematically connecting these areas of research with Dual Inheritance Theory and Cultural Evolution. The present dissertation is an attempt in this direction. A useful starting point is the various biases and learning strategies that have been theoretically and empirically identified (Rendell et al., 2011 offer a catalogue based on their social learning tournament.). These include individual-difference biases like the success bias shown in Chapter 3, frequency dependent biases, like those tested in Chapter 4, individual differences in the application of these biases (IQ was identified as one such individual difference in Chapter 4).

Finally, the Cumulative Cultural Brain Hypothesis predicts innovative ancestors that create knowledge worth exploiting via social learning and an environment where that knowledge translates to survival. We see such innovativeness in our closest cousins (Hopper et al., 2014; Manrique, Völter, & Call, 2013). The ecological prediction is more difficult to predict, but bipedal humans may have had large home ranges from which to forage and animal social learning does seem to focus on food locations and exploitation.

6.2.6 What Are The Other Central Questions?

There are of course many central questions that flow from many of these issues, but whose discussion would lead to a book length discussion section. Here are examples of such central questions:

1. How have human social networks evolved and what role do they play in cultural evolution? Recent research reveals that human social networks are more efficient for information transmission than other primates (Pasquaretta et al., 2014). Chapter 5 is a

first attempt to theoretically understand the origins of these networks and their implications for culture.

2. How do innovations emerge in cultural evolution? Both Charles Darwin and Alfred Wallace arrived upon natural selection at around the same time. Both Isaac Newton and Gottfried Leibniz arrived upon calculus at around the same time. In both these cases, innovation might be seen as cultural recombination of the culture being transmitted at the time. But, it was only Darwin and Wallace who arrived upon natural selection; only Newton and Leibniz who arrived upon calculus. Individual differences matter. In other cases, innovations were entirely serendipitous (e.g. penicillin, vulcanized rubber, microwave heating, Velcro, and Teflon).
3. How do characteristics of the learner, content, and model interact? For example, are individuals more likely to learn some kinds of content from ingroup members than outgroup members?
4. What role has cultural group selection played in the evolution of our species (for review see Chudek et al., 2015; Richerson et al., 2015)?
5. How has cultural evolution changed our biology (both genetically and developmentally)? Some possible examples in recent times include lactase persistence (Laland et al., 2010), intelligence (Cochran, Hardy, & Harpending, 2006), reading ability (McCandliss, Cohen, & Dehaene, 2003), individualism and collectivism (Chiao & Blizinsky, 2009), and ability to use tonal language (Dediu & Ladd, 2007). What are the processes through which this gene-culture coevolution takes place? For example, the line between developmental and genetic changes are blurred if an adaptive trait is acquired via learning, reaches fixation, allowing selection to exert pressures on genes that allow the trait to be acquired more

- quickly – i.e. a Baldwin effect (Burman, 2013). Reading ability is a good candidate for a cultural trait going through such selection. Chapter 2 is a Dual Inheritance Theory model about the relationship between genes (for brains and social learning) and culture.
6. What role have founder effects and bottlenecks played in the evolution of our species and in cultural differences between populations? An extreme example can be found in the Pingelap atoll in the Pacific Ocean. In most places in the world complete colorblindness (achromatopsia) is present in 1 in 30,000. The Pingelapese rate is 1 in 12 (Sacks, 1997). The disease is unlikely to be an adaptation or mistake, but can be traced to a population bottleneck caused by a 1775 typhoon, in which most of the population died (Sundin et al., 2000). In a similar fashion, some of what we consider uniquely human may be accidental characteristics (both detrimental and beneficial) caused by population bottlenecks, at least in the distant past (Hawks, Hunley, Lee, & Wolpoff, 2000). In recent times, population differences may also be a result of small migrant founding populations. Karmin et al. (2015) find a drop in Y-chromosome diversity coinciding with the rise of culture. Such founder effects have been measured both genetically and culturally (at least for language; Atkinson, 2011) As one might predict, there is more genetic and linguistic diversity in Africa than anywhere else and both of these decrease with distance from Africa.
 7. How does cultural content shape other cultural content? How does one innovation open new “thought spaces” and affect the genesis of other innovations? An example of work in this area is Henrich, Boyd, and Richerson’s (2012) research on the evolution of monogamous marriage. In Section 6.3 of the Conclusion, I speculate about the possible role of technology in shaping our theories. Ultimately, this becomes the science of

history and is probably the most neglected area of research in Cultural Evolution, because of how difficult it is to build and test formal models. However, it is also the area that may have the most to say about progress, including scientific progress.

6.3 Technology Shapes Our Theories

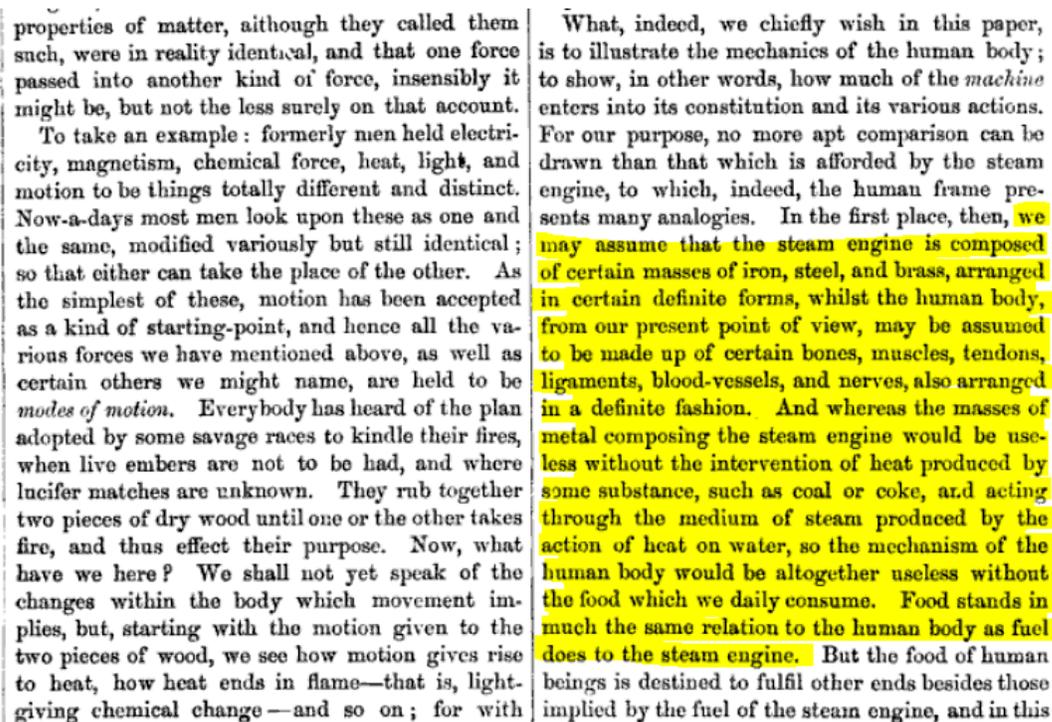
I have used the analogy of hardware and software to describe our brains and culture, respectively. As a software engineer, I'm particularly drawn to these analogies, but such analogies are part of a more challenging and perhaps neglected aspect of cultural evolution; the way cultural content affects cultural content. The way ideas affect other ideas; in this case the way technology can affect our theories and open up new "thought spaces". In the next section, I'll briefly go over the way in which technology shapes our theories, particular those related to our own species. I'll then discuss the process of theory building and testing that I've taken in this dissertation. I will end with a future directions on how technology shapes our theories.

Humans have a tendency to use the latest technology as an analogy for what we consider the greatest technology created by nature – ourselves. The 17th century was the era of sophisticated mechanical devices. In 1600 Galileo published *Le Meccaniche* ("On Mechanics") and by 1642 Blaise Pascal had created the first mechanical calculator. It should come as no surprise that it was between 1641 and 1649 that Descartes wrote several works arguing that the human body is like a machine (controlled by a non-material mind or soul), including *Meditationes de Prima Philosophia* ("Meditations on First Philosophy") in 1641 and *La description du corps humain* ("The Description of the Human Body") in 1647. "And as a clock composed of wheels and counter-weights no less exactly observes the laws of nature when it is badly made, and does not show the time properly, than when it entirely satisfies the wishes of its maker, and as, if I consider the body of a man as being a sort of machine so built up and composed of nerves, muscles, veins, blood and skin..." (Descartes, 1641). Descartes

wasn't alone. Hobbes (1651) goes further "For what is the 'heart' but a 'spring'; and the 'nerves' but so many 'strings'; and the 'joints' but so many 'wheels,' giving motion to the whole body, such as was intended by the artificer?" A century later, James Watts invented an improved steam engine, a key innovation helping to launch the Industrial Revolution. By the 19th century, even in the popular press, the body was likened to a steam engine, as the excerpt from an 1869 People's Magazine article illustrates (Figure 6.1). We see vestiges of this analogy in early 19th century idioms like "blowing off some steam"²³. Up until the 20th century technology shaped the analogies for the human body, but Descartes non-material mind or soul still controlled the thoughts of the day. But by the mid 20th century, the computing revolution begins and we now have a metaphor for the mind. Even today, psychology is rife with computing metaphors, focused on permanent and temporary storage, input and output. The computational model of the mind is slowly changing as the field updates to more modern forms of computing. Today, some evolutionary psychologists who take a modular perspective on the mind use the analogy of an iPhone with apps (Kurzban, 2012). With the advent of quantum computing (my current home Vancouver, is home to one of the first commercial

²³ blow off steam. Source: The American Heritage Dictionary of Idioms by Christine Ammer. (2003, 1997).

quantum computing companies, D-wave), I expect to see more “quantum” analogies of the mind²⁴.



properties of matter, although they called them such, were in reality identical, and that one force passed into another kind of force, insensibly it might be, but not the less surely on that account. To take an example : formerly men held electricity, magnetism, chemical force, heat, light, and motion to be things totally different and distinct. Now-a-days most men look upon these as one and the same, modified variously but still identical ; so that either can take the place of the other. As the simplest of these, motion has been accepted as a kind of starting-point, and hence all the various forces we have mentioned above, as well as certain others we might name, are held to be *modes of motion*. Everybody has heard of the plan adopted by some savage races to kindle their fires, when live embers are not to be had, and where lucifer matches are unknown. They rub together two pieces of dry wood until one or the other takes fire, and thus effect their purpose. Now, what have we here ? We shall not yet speak of the changes within the body which movement implies, but, starting with the motion given to the two pieces of wood, we see how motion gives rise to heat, how heat ends in flame—that is, light-giving chemical change—and so on ; for with

What, indeed, we chiefly wish in this paper, is to illustrate the mechanics of the human body ; to show, in other words, how much of the *machine* enters into its constitution and its various actions. For our purpose, no more apt comparison can be drawn than that which is afforded by the steam engine, to which, indeed, the human frame presents many analogies. In the first place, then, we may assume that the steam engine is composed of certain masses of iron, steel, and brass, arranged in certain definite forms, whilst the human body, from our present point of view, may be assumed to be made up of certain bones, muscles, tendons, ligaments, blood-vessels, and nerves, also arranged in a definite fashion. And whereas the masses of metal composing the steam engine would be useless without the intervention of heat produced by some substance, such as coal or coke, and acting through the medium of steam produced by the action of heat on water, so the mechanism of the human body would be altogether useless without the food which we daily consume. Food stands in much the same relation to the human body as fuel does to the steam engine. But the food of human beings is destined to fulfil other ends besides those implied by the fuel of the stean engine, and in this

Figure 6.1. Excerpt from 1869 People’s Magazine article on Muscular Motion.

Technology shapes our metaphors, analogies and theories (Gigerenzer and Goldstein (1996) call this a tools-to-theories heuristic). I suspect a new generation of “digital natives” more familiar with the interaction between software and hardware, Turing completeness (where the computer can fully represent the computer itself), virtualization and abstraction (with software able to replicate hardware, albeit slower) will find Dual Inheritance Theory far more intuitive. However, recognizing the human tendency to use technology as metaphors should also help us identify the limitations of those metaphors. In the case of Dual Inheritance Theory, a useful metaphor is quickly updating software running on more slowly upgraded hardware, but a key difference is that here, the software

²⁴ Some researchers have already speculated that the mind may in fact be a quantum computer (Koch & Hepp, 2006).

changes the hardware, enabling better software – this is the essence of the model presented in Chapter 2. Nevertheless, I am hopeful that with newer analogies, particularly those emerging from the spectacular advancements in machine learning (e.g. Hinton et al., 2012) we will grow closer to a better description of our species. In the meantime, Dual Inheritance Theory and in particular Cultural Evolution helps us explain how it is that we can even use cultural knowledge to open new “thought spaces” and develop better theories.

References

- Agafonova, N., Aleksandrov, A., Altinok, O., Alvarez Sanchez, P., Aoki, S., Ariga, A., . . . Adam, T. (2012). Measurement of the neutrino velocity with the OPERA detector in the CNGS beam. *JHEP*, *1210*, 093.
- Aiello, L. C., & Wheeler, P. (1995). The expensive-tissue hypothesis: the brain and the digestive system in human and primate evolution. *Current Anthropology*, *36*, 199-221.
- Albarracín, D., & Vargas, P. (2010). Attitudes and persuasion: From biology to social responses to persuasive intent. In S. T. Fiske, D. T. Gilbert & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 394-427). Hoboken, NJ: Wiley.
- Alison, J. (1966). Lemur Social Behavior and Primate Intelligence. *Science*, *153*, 501-506.
- Aoki, K., & Feldman, M. W. (2014). Evolution of learning strategies in temporally and spatially variable environments: a review of theory. *Theoretical Population Biology*, *91*, 3-19.
- Aoki, K., Lehmann, L., & Feldman, M. W. (2011). Rates of cultural change and patterns of cultural accumulation in stochastic models of social transmission. *Theoretical Population Biology*, *79*, 192-202.
- Apicella, C. L., Marlowe, F. W., Fowler, J. H., & Christakis, N. A. (2012). Social networks and cooperation in hunter-gatherers. *Nature*, *481*, 497-U109.
- Aron, A., Aron, E. N., & Smollan, D. (1992). Inclusion of Other in the Self Scale and the structure of interpersonal closeness. *Journal of Personality and Social Psychology*, *63*, 596.
- Asendorpf, J. B., & Wilpers, S. (1998). Personality effects on social relationships. *Journal of Personality and Social Psychology*, *74*, 1531-1544.
- Atkinson, Q. D. (2011). Phonemic diversity supports a serial founder effect model of language expansion from Africa. *Science*, *332*, 346-349.

- Axelrod, R. M. (1997). *The complexity of cooperation: Agent-based models of competition and collaboration*. Princeton, NJ: Princeton University Press.
- Bailey, D. H., & Geary, D. C. (2009). Hominid brain evolution. *Human Nature*, 20, 67-79.
- Balding, D. J., & Nichols, R. A. (1995). A method for quantifying differentiation between populations at multi-allelic loci and its implications for investigating identity and paternity. In B. S. Weir (Ed.), *Human Identification: The Use of DNA Markers* (Vol. 4, pp. 3-12). Dordrecht, Netherlands: Springer Netherlands.
- Bandura, A. (1977). *Social learning theory*. Oxford, England: Prentice-Hall.
- Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of Personality and Social Psychology*, 71, 230.
- Barrett, H. C. (2014). *The Shape of Thought: How Mental Adaptations Evolve*. Oxford University Press.
- Barton, R. A. (1996). Neocortex size and behavioural ecology in primates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 263, 173-177.
- Batchelder, W. H. (1975). Individual differences and the all-or-none vs incremental learning controversy. *Journal of Mathematical Psychology*, 12, 53-74.
- Baum, W. M., Richerson, P. J., Efferson, C. M., & Paciotti, B. M. (2004). Cultural evolution in laboratory microsocieties including traditions of rule giving and rule following. *Evolution and Human Behavior*, 25, 305-326.
- Bell, A. V., Richerson, P. J., & McElreath, R. (2009). Culture rather than genes provides greater scope for the evolution of large-scale human prosociality. *Proceedings of the National Academy of Sciences*, 106, 17671-17674.
- Bem, D. J. (2011). Feeling the future: experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, 100, 407.

- Berger, J. (2013). *Contagious: Why things catch on*. New York, NY: Simon and Schuster.
- Berger, J., & Heath, C. (2007). Where consumers diverge from others: Identity signaling and product domains. *Journal of Consumer Research*, *34*, 121-134.
- Berger, J., & Heath, C. (2008). Who drives divergence? Identity signaling, outgroup dissimilarity, and the abandonment of cultural tastes. *Journal of Personality and Social Psychology*, *95*, 593-607.
- Berger, J., & Schwartz, E. M. (2011). What drives immediate and ongoing word of mouth? *Journal of Marketing Research*, *48*, 869-880.
- Binford, L. R. (2001). Constructing frames of reference: an analytical method for archaeological theory building using ethnographic and environmental data sets. *University of California, Berkeley*.
- Boesch, C. (1991). Teaching among wild chimpanzees. *Animal Behaviour*, *41*, 530-532.
- Bond, R., & Smith, P. B. (1996). Culture and conformity: A meta-analysis of studies using Asch's (1952b, 1956) line judgment task. *Psychological Bulletin*, *119*, 111-137.
- Bourgeois, M. J. (2002). Heritability of attitudes constrains dynamic social impact. *Personality and Social Psychology Bulletin*, *28*, 1063-1072.
- Bower, J. M., & Bolouri, H. (2001). *Computational modeling of genetic and biochemical networks*. Cambridge, MA: MIT press.
- Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. Chicago, IL: University of Chicago Press.
- Boyd, R., & Richerson, P. J. (1988). An evolutionary model of social learning: the effects of spatial and temporal variation. *Social learning: psychological and biological perspectives*, 29-48.
- Boyd, R., & Richerson, P. J. (1996). *Why culture is common, but cultural evolution is rare*. Paper presented at the Proceedings of the British Academy.

- Boyd, R., & Richerson, P. J. (2009). Culture and the evolution of human cooperation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*, 3281-3288.
- Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*, 10918-10925.
- Brent, Lauren J. N., Franks, Daniel W., Foster, Emma A., Balcomb, Kenneth C., Cant, Michael A., & Croft, Darren P. (2015). Ecological Knowledge, Leadership, and the Evolution of Menopause in Killer Whales. *Current Biology*, *25*, 746-750.
- Brosseau-Liard, P. E., & Birch, S. A. (2010). 'I bet you know more and are nicer too!': what children infer from others' accuracy. *Developmental Science*, *13*, 772-778.
- Brown, J. J., & Reingen, P. H. (1987). Social ties and word-of-mouth referral behavior. *Journal of Consumer Research*, *14*, 350-362.
- Buckels, E. E., Beall, A. T., Hofer, M. K., Lin, E. Y., Zhou, Z., & Schaller, M. (2015). Individual Differences in Activation of the Parental Care Motivational System: Assessment, Prediction, and Implications.
- Buenstorf, G., & Cordes, C. (2008). Can sustainable consumption be learned? A model of cultural evolution. *Ecological Economics*, *67*, 646-657.
- Burman, J. T. (2013). Updating the Baldwin effect: The biological levels behind Piaget's new theory. *New Ideas in Psychology*, *31*, 363-373.
- Byrne, D. (1971). *The attraction paradigm*. San Diego, CA: Academic Press.
- Cacioppo, J. T., Petty, R. E., Feinstein, J. A., & Jarvis, W. B. G. (1996). Dispositional differences in cognitive motivation: The life and times of individuals varying in need for cognition. *Psychological Bulletin*, *119*, 197-253.

- Caldwell, C. A., & Millen, A. E. (2010). Human cumulative culture in the laboratory: Effects of (micro) population size. *Learning & behavior*, *38*, 310-318.
- Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in cognitive sciences*, *12*, 187-192.
- Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*: Princeton University Press.
- Chagnon, N. A. (1988). Life histories, blood revenge, and warfare in a tribal population. *Science*, *239*, 985-992.
- Chan, E. (2009). *Quantitative trading: how to build your own algorithmic trading business* (Vol. 430): John Wiley & Sons.
- Chapais, B. (2009). *Primeval kinship: How pair-bonding gave birth to human society*: Harvard University Press.
- Charvet, C. J., & Finlay, B. L. (2012). Embracing covariation in brain evolution: large brains, extended development, and flexible primate social systems. *Progress in brain research*, *195*, 71.
- Chen, S. X., & Bond, M. H. (2010). Two languages, two personalities? Examining language effects on the expression of personality in a bilingual context. *Personality and Social Psychology Bulletin*, *36*, 1514-1528.
- Cheng, J. T., Tracy, J. L., & Henrich, J. (2010). Pride, personality, and the evolutionary foundations of human social status. *Evolution and Human Behavior*, *31*, 334-347.
- Chiao, J. Y., & Blizinsky, K. D. (2009). Culture–gene coevolution of individualism–collectivism and the serotonin transporter gene. *Proceedings of the Royal Society B: Biological Sciences*, rspb20091650.

- Chiu, C.-y., Morris, M. W., Hong, Y.-y., & Menon, T. (2000). Motivated cultural cognition: The impact of implicit cultural theories on dispositional attribution varies as a function of need for closure. *Journal of Personality and Social Psychology*, *78*, 247-259.
- Chua, R. Y.-J., & Morris, M. W. (2006). Dynamics of trust in guanxi networks. In R.-Y. Chen (Ed.), *National culture and groups* (Vol. 9, pp. 95-113). Oxford, United Kingdom: JAI Press.
- Chudek, M., Brosseau-Liard, P. E., Birch, S., & Henrich, J. (2013). Culture-gene coevolutionary theory and children's selective social learning. In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 181). Oxford: Oxford University Press.
- Chudek, M., Heller, S., Birch, S. A., & Henrich, J. (2012). Prestige-biased cultural learning: bystander's differential attention to potential models influences children's learning. *Evolution and Human Behavior*, *33*, 46-56.
- Chudek, M., & Henrich, J. (2011). Culture-gene coevolution, norm-psychology and the emergence of human prosociality. *Trends in cognitive sciences*, *15*, 218-226.
- Chudek, M., Muthukrishna, M., & Henrich, J. (2015). Cultural Evolution. In D. M. Buss (Ed.), *The Handbook of Evolutionary Psychology* (2nd ed., Vol. 2): John Wiley and Sons.
- Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology*, *55*, 591-621.
- Cialdini, R. B., Wosinska, W., Barrett, D. W., Butner, J., & Gornik-Durose, M. (1999). Compliance with a request in two cultures: The differential influence of social proof and commitment/consistency on collectivists and individualists. *Personality and Social Psychology Bulletin*, *25*, 1242-1253.

- Claidière, N., Bowler, M., Brookes, S., Brown, R., & Whiten, A. (2014). Frequency of Behavior Witnessed and Conformity in an Everyday Social Context. *PLoS One*, *9*, e99874.
- Claidière, N., Bowler, M., & Whiten, A. (2012). Evidence for weak or linear conformity but not for hyper-conformity in an everyday social learning context. *PLoS One*, *7*, e30970.
- Claidière, N., & Sperber, D. (2010). Imitation explains the propagation, not the stability of animal culture. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 651-659.
- Claidière, N., & Whiten, A. (2012). Integrating the study of conformity and culture in humans and nonhuman animals. *Psychological Bulletin*, *138*, 126.
- Cochran, G., Hardy, J., & Harpending, H. (2006). Natural history of Ashkenazi intelligence. *Journal of biosocial science*, *38*, 659-693.
- Cohen, J. (1994). The earth is round ($p < .05$). *American psychologist*, *49*, 997-1003.
- Collar, A. (2007). Network theory and religious innovation. *Mediterranean Historical Review*, *22*, 149-162.
- Collard, M., Buchanan, B., Morin, J., & Costopoulos, A. (2011). What drives the evolution of hunter–gatherer subsistence technology? A reanalysis of the risk hypothesis with data from the Pacific Northwest. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 1129-1138.
- Collard, M., Kemery, M., & Banks, S. (2005). Causes of toolkit variation among hunter-gatherers: a test of four competing hypotheses. *Canadian Journal of Archaeology/Journal Canadien d'Archéologie*, 1-19.
- Collard, M., Ruttle, A., Buchanan, B., & O'Brien, M. J. (2012). Risk of Resource Failure and Toolkit Variation in Small-Scale Farmers and Herders. *PLoS One*, *7*.

- Conway, L. G., III, & Schaller, M. (2007). How communication shapes culture. In K. Fiedler (Ed.), *Social communication* (pp. 107-127). New York, NY: Psychology Press.
- Cosmides, L., & Tooby, J. (1992). Cognitive adaptations for social exchange. *The adapted mind*, 163-228.
- Coultas, J. C. (2004). When in Rome... An evolutionary perspective on conformity. *Group Processes & Intergroup Relations*, 7, 317-331.
- Crandall, C. S. (1988). Social contagion of binge eating. *Journal of Personality and Social Psychology*, 55, 588-598.
- Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, 129, 414-446.
- Cullum, J., & Harton, H. C. (2007). Cultural evolution: Interpersonal influence, issue importance, and the development of shared attitudes in college residence halls. *Personality and Social Psychology Bulletin*, 33, 1327-1339.
- Cumming, G. (2013). The new statistics why and how. *Psychological Science*, 0956797613504966.
- Darwin, C. (1859). On the origins of species by means of natural selection. *London: Murray*.
- Dávid-Barrett, T., & Dunbar, R. (2013). Processing power limits social group size: computational evidence for the cognitive costs of sociality. *Proceedings of the Royal Society B: Biological Sciences*, 280, 20131151.
- Dean, L. G., Kendal, R. L., Schapiro, S. J., Thierry, B., & Laland, K. N. (2012). Identification of the social and cognitive processes underlying human cumulative culture. *Science*, 335, 1114-1118.
- Deaner, R. O., Isler, K., Burkart, J., & van Schaik, C. (2007). Overall brain size, and not encephalization quotient, best predicts cognitive ability across non-human primates. *Brain, Behavior and Evolution*, 70, 115-124.

- Dediu, D., & Ladd, D. R. (2007). Linguistic tone is related to the population frequency of the adaptive haplogroups of two brain size genes, ASPM and Microcephalin. *Proceedings of the National Academy of Sciences*, *104*, 10944-10949.
- Descartes, R. (1641). *Meditations*. In E. S. Haldane (Ed.), *The Philosophical Works of Descartes*. Cambridge: Cambridge University Press.
- Dobzhansky, T. (1973). Nothing in biology makes sense except in the light of evolution.
- Doyen, S., Klein, O., Pichon, C.-L., & Cleeremans, A. (2012). Behavioral Priming: It's All in the Mind, but Whose Mind? *PLoS One*, *7*, e29081.
- Dunbar, R. I. (1992). Neocortex size as a constraint on group size in primates. *Journal of Human Evolution*, *22*, 469-493.
- Dunbar, R. I. (2003). The social brain: mind, language, and society in evolutionary perspective. *Annual review of anthropology*, 163-181.
- Dunbar, R. I. (2009). The social brain hypothesis and its implications for social evolution. *Annals of human biology*, *36*, 562-572.
- Dunbar, R. I., & Shultz, S. (2007a). Evolution in the social brain. *Science*, *317*, 1344-1347.
- Dunbar, R. I., & Shultz, S. (2007b). Understanding primate brain evolution. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *362*, 649-658.
- Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, *6*, 178-190.
- Edinborough, K. (2009). Population history, abrupt climate change, and evolution of arrowhead technology in Mesolithic south Scandinavia (pp. 191-202): Berkeley, CA: University of California Press.

- Efferson, C., Lalive, R., Richerson, P. J., McElreath, R., & Lubell, M. (2008). Conformists and mavericks: the empirics of frequency-dependent cultural transmission. *Evolution and Human Behavior, 29*, 56-64.
- Efferson, C., Richerson, P. J., McElreath, R., Lubell, M., Edsten, E., Waring, T. M., . . . Baum, W. (2007). Learning, productivity, and noise: an experimental study of cultural transmission on the Bolivian Altiplano. *Evolution and Human Behavior, 28*, 11-17.
- Eich, E. (2014). Business not as usual. *Psychological Science, 25*, 3-6.
- Emery, N. J. (2006). Cognitive ornithology: the evolution of avian intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 361*, 23-43.
- Emlen, S. T. (1995). An evolutionary theory of the family. *Proceedings of the National Academy of Sciences, 92*, 8092-8099.
- Enquist, M., Strimling, P., Eriksson, K., Laland, K., & Sjostrand, J. (2010). One cultural parent makes no culture. *Animal Behaviour, 79*, 1353-1362.
- Epstein, J. M. (2006). *Generative social science: Studies in agent-based computational modeling*. Princeton, NJ: Princeton University Press.
- Eriksson, K., & Coultas, J. (2009). Are people really conformist-biased? An empirical test and a new mathematical model. *Journal of Evolutionary Psychology, 7*, 5-21.
- Eriksson, K., Enquist, M., & Ghirlanda, S. (2007). Critical points in current theory of conformist social learning. *Journal of Evolutionary Psychology, 5*, 67-87.
- Eubank, S., Guclu, H., Kumar, V. A., Marathe, M. V., Srinivasan, A., Toroczkai, Z., & Wang, N. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature, 429*, 180-184.
- Farnen, R. F., & Meloen, J. (2000). *Democracy, authoritarianism and education: a cross-national empirical survey*. New York, NY: St. Martin's Press.

- Feldman, M. W., & Laland, K. N. (1996). Gene-culture coevolutionary theory. *Trends in Ecology & Evolution*, *11*, 453-457.
- Festinger, L., Schachter, S., & Back, K. (1950). The spatial ecology of group formation. In L. Festinger, K. W. Back & S. Schachter (Eds.), *Social pressures in informal groups* (pp. 33-60). Stanford, CA: Stanford University Press.
- Fiske, S. T., & Taylor, S. E. (2013). *Social cognition: From brains to culture*: Sage.
- Flynn, E., & Smith, K. (2012). Investigating the mechanisms of cultural acquisition: How pervasive is overimitation in adults? *Social Psychology*, *43*, 185.
- Flynn, E., & Whiten, A. (2012). Experimental “Microcultures” in Young Children: Identifying Biographic, Cognitive, and Social Predictors of Information Transmission. *Child development*, *83*, 911-925.
- Fogarty, L., Strimling, P., & Laland, K. N. (2011). The evolution of teaching. *Evolution*, *65*, 2760-2770.
- Foley, R. A., Lee, P. C., Widdowson, E., Knight, C., & Jonxis, J. (1991). Ecology and energetics of encephalization in hominid evolution [and discussion]. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *334*, 223-232.
- Foster, E. A., Franks, D. W., Mazzi, S., Darden, S. K., Balcomb, K. C., Ford, J. K. B., & Croft, D. P. (2012). Adaptive Prolonged Postreproductive Life Span in Killer Whales. *Science*, *337*, 1313.
- Fowler, J. H., Christakis, N. A., Steptoe, & Roux, D. (2009). Dynamic spread of happiness in a large social network: longitudinal analysis of the Framingham Heart Study social network. *BMJ: British medical journal*, 23-27.
- Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic perspectives*, 25-42.

- Gavrilets, S., & Vose, A. (2006). The dynamics of Machiavellian intelligence. *Proceedings of the National Academy of Sciences*, *103*, 16823-16828.
- Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., . . . Arnadottir, J. (2011). Differences between tight and loose cultures: A 33-nation study. *Science*, *332*, 1100-1104.
- Gentilucci, M., & Corballis, M. C. (2006). From manual gesture to speech: a gradual transition. *Neuroscience & Biobehavioral Reviews*, *30*, 949-960.
- Gigerenzer, G., & Goldstein, D. G. (1996). Mind as computer: Birth of a metaphor. *Creativity Research Journal*, *9*, 131-144.
- Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*: Mit Press.
- Gintis, H. (2000). Beyond Homo economicus: evidence from experimental economics. *Ecological Economics*, *35*, 311-322.
- Gintis, H. (2011). Gene–culture coevolution and the nature of human sociality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *366*, 878-888.
- Gupta, A. K., & Nadarajah, S. (2004). *Handbook of beta distribution and its applications*. Boca Raton, FL: CRC Press.
- Gurven, M., Kaplan, H., & Gutierrez, M. (2006). How long does it take to become a proficient hunter? Implications for the evolution of extended development and long life span. *Journal of Human Evolution*, *51*, 454-470.
- Haider, M., & Kreps, G. L. (2004). Forty years of diffusion of innovations: utility and value in public health. *Journal of health communication*, *9*, 3-11.
- Harihara, M. (2014). *Cultural differences in social network structures: Comparative study in the United States, Japan, and Korea*. Paper presented at the Fifteenth Annual Meeting for the Society of Personality and Social Psychology, Austin, TX.

- Harlow, L. L., Mulaik, S. A., & Steiger, J. H. (1997). *What if there were no significance tests?* : Psychology Press.
- Harmand, S., Lewis, J. E., Feibel, C. S., Lepre, C. J., Prat, S., Lenoble, A., . . . Roche, H. (2015). 3.3-million-year-old stone tools from Lomekwi 3, West Turkana, Kenya. *Nature*, *521*, 310-315.
- Harton, H. C., & Bourgeois, M. J. (2003). Cultural elements emerge from dynamic social impact. In M. Schaller & C. S. Crandall (Eds.), *The psychological foundations of culture* (pp. 41-75). New York, NY: Psychology Press.
- Harton, H. C., & Bullock, M. (2007). Dynamic social impact: A theory of the origins and evolution of culture. *Social and Personality Psychology Compass*, *1*, 521-540.
- Hastie, R., & Kameda, T. (2005). The Robust Beauty of Majority Rules in Group Decisions. *Psychological Review*, *112*, 494-508.
- Hastie, R., & Stasser, G. (2000). Computer simulation methods for social psychology. *Handbook of research methods in social and personality psychology*, 85-114.
- Haun, Daniel B. M., Rekers, Y., & Tomasello, M. (2012). Majority-Biased Transmission in Chimpanzees and Human Children, but Not Orangutans. *Current Biology*, *22*, 727-731.
- Hawks, J., Hunley, K., Lee, S.-H., & Wolpoff, M. (2000). Population Bottlenecks and Pleistocene Human Evolution. *Molecular Biology and Evolution*, *17*, 2-22.
- Heine, S. J., & Buchtel, E. E. (2009). Personality: The universal and the culturally specific. *Annual Review of psychology*, *60*, 369-394.
- Heine, S. J., & Norenzayan, A. (2006). Toward a psychological science for a cultural species. *Perspectives on Psychological Science*, *1*, 251-269.
- Henneberg, M. (1988). Decrease of human skull size in the Holocene. *Human Biology*, 395-405.

- Henrich, J. (2004a). Cultural group selection, coevolutionary processes and large-scale cooperation. *Journal of Economic Behavior & Organization*, 53, 3-35.
- Henrich, J. (2004b). Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses: the Tasmanian case. *American Antiquity*, 197-214.
- Henrich, J. (2009a). The evolution of costly displays, cooperation and religion: Credibility enhancing displays and their implications for cultural evolution. *Evolution and Human Behavior*, 30, 244-260.
- Henrich, J. (2012). Too late: models of cultural evolution and group selection have already proved useful. *The False Allure of Group Selection*. Retrieved May 25, 2015, from <http://edge.org/conversation/the-false-allure-of-group-selection>
- Henrich, J. (forthcoming). *The secret of our success: How learning from others drove human evolution, domesticated our species, and made us smart*. Princeton, NJ: Princeton University Press.
- Henrich, J. (Ed.). (2009b). *The evolution of innovation-enhancing institutions*. Cambridge: MIT Press.
- Henrich, J., & Boyd, R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, 19, 215-241.
- Henrich, J., & Boyd, R. (2002). On modeling cognition and culture. *Journal of Cognition and Culture*, 2, 87-112.
- Henrich, J., & Boyd, R. (2008). Division of labor, economic specialization, and the evolution of social stratification. *Current Anthropology*, 49, 715-724.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American Economic Review*, 73-78.

- Henrich, J., Boyd, R., & Richerson, P. J. (2008). Five misunderstandings about cultural evolution. *Human Nature, 19*, 119-137.
- Henrich, J., Boyd, R., & Richerson, P. J. (2012). The puzzle of monogamous marriage. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*, 657-669.
- Henrich, J., & Broesch, J. (2011). On the nature of cultural transmission networks: evidence from Fijian villages for adaptive learning biases. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*, 1139-1148.
- Henrich, J., & Gil-White, F. J. (2001). The evolution of prestige: Freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior, 22*, 165-196.
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences, 33*, 61-83.
- Henrich, J., & Henrich, N. (2010). The evolution of cultural adaptations: Fijian food taboos protect against dangerous marine toxins. *Proceedings of the Royal Society B: Biological Sciences, 277*, 3715-3724.
- Henrich, J., & McElreath, R. (2003). The evolution of cultural evolution. *Evolutionary Anthropology: Issues, News, and Reviews, 12*, 123-135.
- Herrmann, E., Call, J., Hernández-Lloreda, M. V., Hare, B., & Tomasello, M. (2007). Humans have evolved specialized skills of social cognition: the cultural intelligence hypothesis. *Science, 317*, 1360-1366.
- Herrmann, P. A., Legare, C. H., Harris, P. L., & Whitehouse, H. (2013). Stick to the script: The effect of witnessing multiple actors on children's imitation. *Cognition, 129*, 536-543.

- Herrmann, P. A., Legare, C. H., Harris, P. L., & Whitehouse, H. (in press). Stick to the script: The effect of witnessing multiple actors on children's imitation. *Cognition*.
- Heyes, C. (2012). Grist and mills: on the cultural origins of cultural learning. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2181-2191.
- Hill, K., & Hurtado, A. M. (2009). *Cooperative breeding in South American hunter-gatherers* (Vol. 276).
- Hill, K. R., Walker, R. S., Božičević, M., Eder, J., Headland, T., Hewlett, B., . . . Wood, B. (2011). Co-residence patterns in hunter-gatherer societies show unique human social structure. *Science*, 331, 1286-1289.
- Hinton, G., Deng, L., Yu, D., Dahl, G. E., Mohamed, A.-r., Jaitly, N., . . . Sainath, T. N. (2012). Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *Signal Processing Magazine, IEEE*, 29, 82-97.
- Hobbes, T. (1651). *Leviathan: Of Man, Being the First Part of Leviathan*. In C. W. Eliot (Ed.), *The Harvard classics*. New York: P.F. Collier & Son.
- Hofstede, G., & McCrae, R. R. (2004). Personality and culture revisited: Linking traits and dimensions of culture. *Cross-cultural research*, 38, 52-88.
- Hofstede, G. H. (2003). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations* (2nd ed.). New York, NY: Sage Publications.
- Hogg, M. A. (2013). Intergroup relations *Handbook of social psychology* (pp. 533-561): Springer.
- Hopper, L. M., Price, S. A., Freeman, H. D., Lambeth, S. P., Schapiro, S. J., & Kendal, R. L. (2014). Influence of personality, age, sex, and estrous state on chimpanzee problem-solving success. *Animal cognition*, 17, 835-847.
- Hoppitt, W., & Laland, K. N. (2013). *Social Learning: An Introduction to Mechanisms, Methods, and Models*: Princeton University Press.

- Hornby, G. S., Globus, A., Linden, D. S., & Lohn, J. D. (2006). *Automated antenna design with evolutionary algorithms*. Paper presented at the AIAA Space.
- Horner, V., & Whiten, A. (2005). Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal cognition*, 8, 164-181.
- Hrdy, S. B. (2009). *Mothers and others: the evolutionary origins of mutual understanding*: Harvard University Press.
- Humphrey, N. K. (1976). The social function of intellect. In P. P. G. Bateson & R. A. Hinde (Eds.), *Growing points in ethology* (pp. 303-317). Cambridge, UK: Cambridge University Press.
- Inoue, S., & Matsuzawa, T. (2007). Working memory of numerals in chimpanzees. *Current Biology*, 17, R1004-R1005.
- Insko, C. A., Thibaut, J. W., Moehle, D., Wilson, M., Diamond, W. D., Gilmore, R., . . . Lipsitz, A. (1980). Social evolution and the emergence of leadership. *Journal of Personality and Social Psychology*, 39, 431-448.
- Isler, K., & Van Schaik, C. P. (2006). Metabolic costs of brain size evolution. *Biology letters*, 2, 557-560.
- Isler, K., & van Schaik, C. P. (2009). The expensive brain: a framework for explaining evolutionary changes in brain size. *Journal of Human Evolution*, 57, 392-400.
- Jackson, M. O. (2010). *Social and economic networks*. Princeton, NJ: Princeton University Press.
- Joffe, T. H. (1997). Social pressures have selected for an extended juvenile period in primates. *Journal of Human Evolution*, 32, 593-605.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). The big five inventory—versions 4a and 54. *Berkeley: University of California, Berkeley, Institute of Personality and Social Research*.

- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative big five trait taxonomy. *Handbook of personality: Theory and research*, 3, 114-158.
- Johnson, S. (2001). *Emergence: The connected lives of ants, brains, cities, and software*. New York, NY: Simon and Schuster.
- Kahneman, D. (2011). *Thinking, fast and slow*: Macmillan.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*, 263-291.
- Kalish, Y., & Robins, G. (2006). Psychological predispositions and network structure: The relationship between individual predispositions, structural holes and network closure. *Social Networks*, 28, 56-84.
- Kameda, T., & Nakanishi, D. (2002). Cost-benefit analysis of social/cultural learning in a nonstationary uncertain environment: An evolutionary simulation and an experiment with human subjects. *Evolution and Human Behavior*, 23, 373-393.
- Kameda, T., Takezawa, M., & Hastie, R. (2003). The logic of social sharing: An evolutionary game analysis of adaptive norm development. *Personality and Social Psychology Review*, 7, 2-19.
- Kao, A. B., & Couzin, I. D. (2014). Decision accuracy in complex environments is often maximized by small group sizes. *Proceedings of the Royal Society B: Biological Sciences*, 281, 20133305.
- Kaplan, H., Gurven, M., Hill, K., & Hurtado, A. M. (2005). The natural history of human food sharing and cooperation: a review and a new multi-individual approach to the negotiation of norms. *Moral sentiments and material interests: The foundations of cooperation in economic life*, 75-113.
- Kaplan, H., Hill, K., Lancaster, J., & Hurtado, A. M. (2000). A theory of human life history evolution: diet, intelligence, and longevity. *Evolutionary Anthropology Issues News and Reviews*, 9, 156-185.

- Karmin, M., Saag, L., Vicente, M., Sayres, M. A. W., Järve, M., Talas, U. G., . . . Mitt, M. (2015). A recent bottleneck of Y chromosome diversity coincides with a global change in culture. *Genome research*, 25, 459-466.
- Kashima, Y., Wilson, S., Lusher, D., Pearson, L. J., & Pearson, C. (2013). The acquisition of perceived descriptive norms as social category learning in social networks. *Social Networks*, 35, 711-719.
- Kendal, J., Giraldeau, L.-A., & Laland, K. (2009). The evolution of social learning rules: payoff-biased and frequency-dependent biased transmission. *Journal of Theoretical Biology*, 260, 210-219.
- Kenney, P. J., & Rice, T. W. (1994). The psychology of political momentum. *Political Research Quarterly*, 47, 923-938.
- Kenrick, D. T., Li, N. P., & Butner, J. (2003). Dynamical evolutionary psychology: Individual decision rules and emergent social norms. *Psychological Review*, 110, 3-28.
- Kerr, N. L. (1998). HARKing: Hypothesizing after the results are known. *Personality and Social Psychology Review*, 2, 196-217.
- Kerr, N. L., & Tindale, R. S. (2004). Group performance and decision making. *Annual Review of Psychology*, 55, 623-655.
- Kim, H., & Markus, H. R. (1999). Deviance or uniqueness, harmony or conformity? A cultural analysis. *Journal of Personality and Social Psychology*, 77, 785-800.
- Kimbrough, E. O., & Vostroknutov, A. (2013). Norms Make Preferences Social *Discussion Papers dp13-01*: Department of Economics, Simon Fraser University.
- King, A. J., & Cowlishaw, G. (2007). When to use social information: the advantage of large group size in individual decision making. *Biology letters*, 3, 137-139.

- Kitano, H. (2002). Computational systems biology. *Nature*, *420*, 206-210.
- Kitayama, S., & Cohen, D. (2010). *Handbook of cultural psychology*: Guilford Press.
- Kline, M. A. (2014). How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behavioral and Brain Sciences, FirstView*, 1-70.
- Kline, M. A., & Boyd, R. (2010). Population size predicts technological complexity in Oceania. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 2559-2564.
- Kobayashi, Y., & Aoki, K. (2012). Innovativeness, population size and cumulative cultural evolution. *Theoretical Population Biology*, *82*, 38-47.
- Koch, C., & Hepp, K. (2006). Quantum mechanics in the brain. *Nature*, *440*, 611-611.
- Koller, M., & Stahel, W. A. (2011). Sharpening wald-type inference in robust regression for small samples. *Computational Statistics & Data Analysis*, *55*, 2504-2515.
- Kotrschal, A., Rogell, B., Bundsen, A., Svensson, B., Zajitschek, S., Brännström, I., . . . Kolm, N. (2013). Artificial selection on relative brain size in the guppy reveals costs and benefits of evolving a larger brain. *Current Biology*, *23*, 168-171.
- Kruglanski, A. W., Webster, D. M., & Klem, A. (1993). Motivated resistance and openness to persuasion in the presence or absence of prior information. *Journal of Personality and Social Psychology*, *65*, 861-876.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*: University of Chicago press.
- Kumkale, G. T., & Albarracín, D. (2004). The Sleeper Effect in Persuasion: A Meta-Analytic Review. *Psychological Bulletin*, *130*, 143-172.
- Kurzban, R. (2012). *Why everyone (else) is a hypocrite: Evolution and the modular mind*: Princeton University Press.

- Laar, C. V., Levin, S., Sinclair, S., & Sidanius, J. (2005). The effect of university roommate contact on ethnic attitudes and behavior. *Journal of Experimental Social Psychology, 41*, 329-345.
- Laland, K. N. (2004). Social learning strategies. *Animal Learning & Behavior, 32*, 4-14.
- Laland, K. N. (2008). Animal cultures. *Current Biology, 18*, R366-R370.
- Laland, K. N., Atton, N., & Webster, M. M. (2011). From fish to fashion: experimental and theoretical insights into the evolution of culture. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*, 958-968.
- Laland, K. N., & Brown, G. (2011). *Sense and nonsense: Evolutionary perspectives on human behaviour*. Oxford University Press.
- Laland, K. N., Odling-Smee, J., & Myles, S. (2010). How culture shaped the human genome: bringing genetics and the human sciences together. *Nature Reviews Genetics, 11*, 137-148.
- Latané, B. (1996). Dynamic social impact: The creation of culture by communication. *Journal of Communication, 46*, 13-25.
- Latané, B., & Bourgeois, M. J. (2001). Successfully simulating dynamic social impact. In J. P. Forgas & K. D. Williams (Eds.), *Social influence: Direct and indirect processes* (Vol. 3, pp. 61-76): Psychology Press.
- Latané, B., & Bourgeois, M. J. (1996). Experimental evidence for dynamic social impact: The emergence of subcultures in electronic groups. *Journal of Communication, 46*, 35-47.
- Latané, B., Liu, J. H., Nowak, A., Bonevento, M., & Zheng, L. (1995). Distance matters: Physical space and social impact. *Personality and Social Psychology Bulletin, 21*, 795-805.
- Latané, B., & Nowak, A. (1994). Attitudes as catastrophes: From dimensions to categories with increasing involvement. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 219-249). San Diego, CA, US: Academic Press.

- Lefebvre, L. (2013). Brains, innovations, tools and cultural transmission in birds, non-human primates, and fossil hominins. *Frontiers in human neuroscience*, 7.
- Legare, C. H., & Clegg, J. M. (2015). The Development of Children's Causal Explanations. In S. Robson & S. Quinn (Eds.), *Routledge International Handbook on Young Children's Thinking and Understanding*: Routledge.
- Lehmann, L., Aoki, K., & Feldman, M. W. (2011). On the number of independent cultural traits carried by individuals and populations. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 424-435.
- Lewis, H. M., & Laland, K. N. (2012). Transmission fidelity is the key to the build-up of cumulative culture. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367, 2171-2180.
- Lind, J., & Lindenfors, P. (2010). The number of cultural traits is correlated with female group size but not with male group size in chimpanzee communities. *PLoS One*, 5, e9241.
- Liu, C., & Wechsler, H. (2002). Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *Image processing, IEEE Transactions on*, 11, 467-476.
- Lumsden, C. J., & Wilson, E. O. (1981). *Genes, mind, and culture: The coevolutionary process*. Cambridge, MA: Harvard University Press.
- Lyons, D. E., Young, A. G., & Keil, F. C. (2007). The hidden structure of overimitation. *Proceedings of the National Academy of Sciences*, 104, 19751-19756.
- MacCoun, R. J. (2012). The burden of social proof: Shared thresholds and social influence. *Psychological Review*, 119, 345.
- Mace, R., & Sear, R. (2005). Are humans cooperative breeders? In E. Voland, A. Chasiotis & W. Schiefenhövel (Eds.), *Grandmotherhood: the evolutionary significance of the second half of female life* (pp. 143–159). New Brunswick, NJ: Rutgers University Press.

- MacLean, E. L., Hare, B., Nunn, C. L., Addessi, E., Amici, F., Anderson, R. C., . . . Zhao, Y. (2014). The evolution of self-control. *Proceedings of the National Academy of Sciences of the United States of America*, *111*, E2140-E2148.
- Mangel, M., & Clark, C. W. (1988). *Dynamic modeling in behavioral ecology*. Princeton, NJ: Princeton University Press.
- Manrique, H. M., Völter, C. J., & Call, J. (2013). Repeated innovation in great apes. *Animal Behaviour*, *85*, 195-202.
- Marquet, P. A., Santoro, C. M., Latorre, C., Standen, V. G., Abades, S. R., Rivadeneira, M. M., . . . Hochberg, M. E. (2012). Emergence of social complexity among coastal hunter-gatherers in the Atacama Desert of northern Chile. *Proceedings of the National Academy of Sciences*, *109*, 14754-14760.
- Martin, C. F., Bhui, R., Bossaerts, P., Matsuzawa, T., & Camerer, C. (2014). Chimpanzee choice rates in competitive games match equilibrium game theory predictions. *Scientific reports*, *4*.
- Martrat, B., Grimalt, J. O., Shackleton, N. J., de Abreu, L., Hutterli, M. A., & Stocker, T. F. (2007). Four Climate Cycles of Recurring Deep and Surface Water Destabilizations on the Iberian Margin. *Science*, *317*, 502-507.
- Mason, W. A., Conrey, F. R., & Smith, E. R. (2007). Situating social influence processes: Dynamic, multidirectional flows of influence within social networks. *Personality and Social Psychology Review*, *11*, 279-300.
- Mathews, T., Hamilton, B. E., & National Center for Health Statistics. (2009). Delayed childbearing: more women are having their first child later in life.

- Matsumoto, D., Yoo, S. H., & Fontaine, J. (2008). Mapping expressive differences around the world: the relationship between emotional display rules and individualism versus collectivism. *Journal of Cross-Cultural Psychology, 39*, 55-74.
- McCandliss, B. D., Cohen, L., & Dehaene, S. (2003). The visual word form area: expertise for reading in the fusiform gyrus. *Trends in cognitive sciences, 7*, 293-299.
- McCrae, R. R. (2002). NEO-PI-R data from 36 cultures. In R. R. McCrae & J. Allik (Eds.), *The five-factor model of personality across cultures* (pp. 105-125). New York, NY: Springer.
- McCrae, R. R., Terracciano, A., & 79 Members of the Personality Profiles of Cultures Project. (2005). Personality profiles of cultures: aggregate personality traits. *Journal of Personality and Social Psychology, 89*, 407.
- McElreath, R., Bell, A. V., Efferson, C., Lubell, M., Richerson, P. J., & Waring, T. (2008). Beyond existence and aiming outside the laboratory: estimating frequency-dependent and pay-off-biased social learning strategies. *Philosophical Transactions of the Royal Society B: Biological Sciences, 363*, 3515-3528.
- McElreath, R., Lubell, M., Richerson, P. J., Waring, T. M., Baum, W., Edsten, E., . . . Paciotti, B. (2005). Applying evolutionary models to the laboratory study of social learning. *Evolution and Human Behavior, 26*, 483-508.
- McGuigan, N., Makinson, J., & Whiten, A. (2011). From over-imitation to super-copying: Adults imitate causally irrelevant aspects of tool use with higher fidelity than young children. *British Journal of Psychology, 102*, 1-18.
- McGuigan, N., Whiten, A., Flynn, E., & Horner, V. (2007). Imitation of causally opaque versus causally transparent tool use by 3-and 5-year-old children. *Cognitive Development, 22*, 353-364.

- Mesoudi, A. (2009). How cultural evolutionary theory can inform social psychology and vice versa. *Psychological Review*, 116, 929-952.
- Mesoudi, A. (2011). An experimental comparison of human social learning strategies: payoff-biased social learning is adaptive but underused. *Evolution and Human Behavior*, 32, 334-342.
- Mesoudi, A., Chang, L., Murray, K., & Lu, H. J. (2015). Higher frequency of social learning in China than in the West shows cultural variation in the dynamics of cultural evolution. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20142209.
- Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 3489-3501.
- Miller, G. (2011). *The mating mind: How sexual choice shaped the evolution of human nature*: Anchor.
- Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81.
- Mischel, W. (2009). The toothbrush problem. *Association for Psychological Science Observer*, 21.
- Moll, H., & Tomasello, M. (2007). *Cooperation and human cognition: the Vygotskian intelligence hypothesis* (Vol. 362).
- Monroe, B. M., & Read, S. J. (2008). A general connectionist model of attitude structure and change: The ACS (Attitudes as Constraint Satisfaction) model. *Psychological Review*, 115, 733-759.
- Morgan, T., Rendell, L., Ehn, M., Hoppitt, W., & Laland, K. (2012). The evolutionary basis of human social learning. *Proceedings of the Royal Society B: Biological Sciences*, 279, 653-662.
- Morgan, T. J. H., Laland, K. N., & Harris, P. L. (2014). The development of adaptive conformity in young children: effects of uncertainty and consensus. *Developmental Science*.

- Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., . . . Laland, K. N. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat Commun*, *6*.
- Moscovici, S. (1980). Toward a theory of conversion behavior. *Advances in experimental social psychology*, *13*, 209-239.
- MSC Software. (2004). Dytran. Retrieved May 25, 2015, from <http://www.mssoftware.com/product/dytran>
- Muthukrishna, M., Shulman, B. W., Vasilescu, V., & Henrich, J. (2013). Sociality influences cultural complexity. *Proceedings of the Royal Society B: Biological Sciences*, *281*, 20132511.
- Nadeau, R., Cloutier, E., & Guay, J.-H. (1993). New evidence about the existence of a bandwagon effect in the opinion formation process. *International Political Science Review*, *14*, 203-213.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R² from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, *4*, 133-142.
- Nakahashi, W., Wakano, J. Y., & Henrich, J. (2012). Adaptive social learning strategies in temporally and spatially varying environments. *Human Nature*, *23*, 386-418.
- Nielsen, M. (2012). Imitation, pretend play, and childhood: Essential elements in the evolution of human culture? *Journal of Comparative Psychology*, *126*, 170.
- Nielsen, M., Subiaul, F., Galef, B., Zentall, T., & Whiten, A. (2012). Social learning in humans and nonhuman animals: theoretical and empirical dissections. *Journal of Comparative Psychology*, *126*, 109.
- Norenzayan, A., & Heine, S. J. (2005). Psychological universals: What are they and how can we know? *Psychological Bulletin*, *131*, 763.
- Nosek, B. A., & Lakens, D. (2014). Registered reports. *Social Psychology*, *45*, 137-141.

- Nowak, A. (2004). Dynamical minimalism: Why less is more in psychology. *Personality and Social Psychology Review*, 8, 183-192.
- Nowak, A., & Latané, B. (1994). Simulating the emergence of social order from individual behavior. *Simulating societies: The computer simulation of social processes*, 63-84.
- Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, 97, 362-376.
- Odling-Smee, F. J., Laland, K. N., & Feldman, M. W. (2003). *Niche construction: the neglected process in evolution*. Princeton University Press.
- Ohtomo, Y., Kakegawa, T., Ishida, A., Nagase, T., & Rosing, M. T. (2014). Evidence for biogenic graphite in early Archaean Isua metasedimentary rocks. *Nature Geosci*, 7, 25-28.
- Oishi, S. (2014). Socioecological psychology. *Annual Review of Psychology*, 65, 581-609.
- Ostrom, T. M. (1988). Computer simulation: The third symbol system. *Journal of Experimental Social Psychology*, 24, 381-392.
- Over, H., & Carpenter, M. (2013). The social side of imitation. *Child Development Perspectives*, 7, 6-11.
- Pagel, M. (2012). *Wired for culture: origins of the human social mind*. WW Norton & Company Incorporated.
- Pasquaretta, C., Leve, M., Claidiere, N., van de Waal, E., Whiten, A., MacIntosh, A. J. J., . . . Sueur, C. (2014). Social networks in primates: smart and tolerant species have more efficient networks. *Scientific reports*, 4.
- Paulhus, D. L., & Trapnell, P. D. (1998). Typological measures of shyness: Additive, interactive, and categorical. *Journal of Research in Personality*, 32, 183-201.
- Pérez-Barbería, F. J., Shultz, S., & Dunbar, R. I. (2007). Evidence for coevolution of sociality and relative brain size in three orders of mammals. *Evolution*, 61, 2811-2821.

- Perreault, C., Moya, C., & Boyd, R. (2012). A Bayesian approach to the evolution of social learning. *Evolution and Human Behavior, 33*, 449-459.
- Persky, J. (1995). Retrospectives: the ethology of homo economicus. *The journal of economic perspectives, 221-231*.
- Petty, R., & Briñol, P. (2011). The elaboration likelihood model. *Handbook of theories of social psychology, 224-245*.
- Pfau, J., Kirley, M., & Kashima, Y. (2013). The co-evolution of cultures, social network communities, and agent locations in an extension of Axelrod's model of cultural dissemination. *Physica A: Statistical Mechanics and its Applications, 392*, 381-391.
- Pianka, E. R. (2011). *Evolutionary ecology*: Eric R. Pianka.
- Pike, T. W., & Laland, K. N. (2010). Conformist learning in nine-spined sticklebacks' foraging decisions. *Biology letters, rsbl20091014*.
- Pollet, T. V., Roberts, S. G., & Dunbar, R. I. (2011). Extraverts have larger social network layers. *Journal of Individual Differences, 32*, 161-169.
- Popper, K. R. (1962). *Conjectures and refutations*. New York: Basic Books.
- Powell, A., Shennan, S., & Thomas, M. G. (2009). Late Pleistocene demography and the appearance of modern human behavior. *Science, 324*, 1298-1301.
- Pradhan, G. R., Tennie, C., & van Schaik, C. P. (2012). Social organization and the evolution of cumulative technology in apes and hominins. *Journal of Human Evolution, 2012*, 1-11.
- Premo, L. S., & Kuhn, S. L. (2010). Modeling effects of local extinctions on culture change and diversity in the Paleolithic. *PLoS One, 5*, e15582.

- Ramírez-Esparza, N., Gosling, S. D., Benet-Martínez, V., Potter, J. P., & Pennebaker, J. W. (2006). Do bilinguals have two personalities? A special case of cultural frame switching. *Journal of Research in Personality, 40*, 99-120.
- Raven, J. C., & Court, J. H. (1998). *Raven's progressive matrices and vocabulary scales*: Oxford Psychologists Press.
- Reader, S. M., Hager, Y., & Laland, K. N. (2011). The evolution of primate general and cultural intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences, 366*, 1017-1027.
- Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences, 99*, 4436-4441.
- Rendell, L., Boyd, R., Cownden, D., Enquist, M., Eriksson, K., Feldman, M. W., . . . Laland, K. N. (2010). Why copy others? Insights from the social learning strategies tournament. *Science, 328*, 208-213.
- Rendell, L., Fogarty, L., Hoppitt, W. J., Morgan, T. J., Webster, M. M., & Laland, K. N. (2011). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends in cognitive sciences, 15*, 68-76.
- Richerson, P., Baldini, R., Bell, A., Demps, K., Frost, K., Hillis, V., . . . Newson, L. (2015). Cultural Group Selection Plays an Essential Role in Explaining Human Cooperation: A Sketch of the Evidence. *Behavioral and Brain Sciences, 1*-71.
- Rogers, A. R. (1988). Does biology constrain culture? *American Anthropologist, 90*, 819-831.
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). New York, NY: Free Press.
- Roth, G., & Dicke, U. (2005). Evolution of the brain and intelligence. *Trends in cognitive sciences, 9*, 250-257.

- Rozeboom, W. W. (1997). Good science is abductive, not hypothetico-deductive. In L. L. Harlow, S. A. Mulaik & J. H. Steiger (Eds.), *What if there were no significance tests* (pp. 335-392).
- Russell, D. A., & Séguin, R. (1982). Reconstructions of the small Cretaceous theropod, *Stenonychosaurus inequalis*, and a hypothetical dinosaurid. *Syllogens*, 37.
- Ryder, A. G., Alden, L. E., & Paulhus, D. L. (2000). Is acculturation unidimensional or bidimensional? A head-to-head comparison in the prediction of personality, self-identity, and adjustment. *Journal of Personality and Social Psychology*, 79, 49.
- Sacks, O. (1997). *The Island of the Colorblind*: New York: AA Knopf.
- Scally, A., & Durbin, R. (2012). Revising the human mutation rate: implications for understanding human evolution. *Nature Reviews Genetics*, 13, 745-753.
- Schaller, M., Conway Iii, L. G., & Tanchuk, T. L. (2002). Selective pressures on the once and future contents of ethnic stereotypes: Effects of the communicability of traits. *Journal of Personality and Social Psychology*, 82, 861-877.
- Schaller, M., & Crandall, C. S. (2003). *The psychological foundations of culture*: Psychology Press.
- Schaller, M., & Murray, D. R. (2008). Pathogens, personality, and culture: Disease prevalence predicts worldwide variability in sociosexuality, extraversion, and openness to experience. *Journal of Personality and Social Psychology*, 95, 212-221.
- Schaller, M., & Murray, D. R. (2011). Infectious disease and the creation of culture. *Advances in culture and psychology*, 1, 99-151.
- Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can there ever be too many options? A meta-analytic review of choice overload. *Journal of Consumer Research*, 37, 409-425.

- Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martínez, V. (2007). The geographic distribution of Big Five personality traits patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology, 38*, 173-212.
- Schnettler, S. (2009). A structured overview of 50 years of small-world research. *Social Networks, 31*, 165-178.
- Schoenemann, P. T. (2006). Evolution of the size and functional areas of the human brain. *Annual review of anthropology, 35*, 379-406.
- Schug, J., Yuki, M., Horikawa, H., & Takemura, K. (2009). Similarity attraction and actually selecting similar others: How cross-societal differences in relational mobility affect interpersonal similarity in Japan and the USA. *Asian Journal of Social Psychology, 12*, 95-103.
- Schwartz, B., & Kliban, K. (2004). *The paradox of choice: Why more is less*. New York: Ecco.
- Selfhout, M., Burk, W., Branje, S., Denissen, J., Van Aken, M., & Meeus, W. (2010). Emerging late adolescent friendship networks and Big Five personality traits: A social network approach. *Journal of Personality, 78*, 509-538.
- Semaw, S., Rogers, M. J., Quade, J., Renne, P. R., Butler, R. F., Dominguez-Rodrigo, M., . . . Simpson, S. W. (2003). 2.6-Million-year-old stone tools and associated bones from OGS-6 and OGS-7, Gona, Afar, Ethiopia. *Journal of Human Evolution, 45*, 169-177.
- Shampanier, K., Mazar, N., & Ariely, D. (2007). Zero as a special price: The true value of free products. *Marketing Science, 26*, 742-757.
- Shultz, S., & Dunbar, R. (2010a). Encephalization is not a universal macroevolutionary phenomenon in mammals but is associated with sociality. *Proceedings of the National Academy of Sciences, 107*, 21582-21586.

- Shultz, S., & Dunbar, R. I. (2010b). Social bonds in birds are associated with brain size and contingent on the correlated evolution of life-history and increased parental investment. *Biological Journal of the Linnean Society*, *100*, 111-123.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 0956797611417632.
- Simonsohn, U. (2013). Just Post It The Lesson From Two Cases of Fabricated Data Detected by Statistics Alone. *Psychological Science*, 0956797613480366.
- Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534.
- Smaldino, P. E. (2013). The cultural evolution of emergent group-level traits. *Behavioral and Brain Sciences*.
- Sol, D., Bacher, S., Reader, S. M., & Lefebvre, L. (2008). Brain size predicts the success of mammal species introduced into novel environments. *the american naturalist*, *172*, S63-S71.
- Striedter, G. F. (2005). *Principles of brain evolution*: Sinauer Associates.
- Sundin, O. H., Yang, J.-M., Li, Y., Zhu, D., Hurd, J. N., Mitchell, T. N., . . . Maumenee, I. H. (2000). Genetic basis of total colourblindness among the Pingelapese islanders. *Nature genetics*, *25*, 289-293.
- Szabó, G., & Tóke, C. (1998). Evolutionary prisoner's dilemma game on a square lattice. *Physical Review E*, *58*, 69.
- Tagliatalata, J. P., Reamer, L., Schapiro, S. J., & Hopkins, W. D. (2012). Social learning of a communicative signal in captive chimpanzees. *Biology letters*, *8*, 498-501.

- Talhelm, T., Zhang, X., Oishi, S., Shimin, C., Duan, D., Lan, X., & Kitayama, S. (2014). Large-scale psychological differences within China explained by rice versus wheat agriculture. *Science*, *344*, 603-608.
- Tanford, S., & Penrod, S. (1983). Computer modeling of influence in the jury: The role of the consistent juror. *Social Psychology Quarterly*, *46*, 200-212.
- Tanford, S., & Penrod, S. (1984). Social Influence Model: A formal integration of research on majority and minority influence processes. *Psychological Bulletin*, *95*, 189-225.
- Tesser, A., & Achee, J. (1994). Aggression, love, conformity, and other social psychological catastrophes. In R. R. Vallacher & A. Nowak (Eds.), *Dynamical systems in social psychology* (pp. 139-167). San Diego, CA: Academic Press.
- Thaler, R. H. (2000). From homo economicus to homo sapiens. *The journal of economic perspectives*, 133-141.
- The Gimp Team. (2012). GIMP. 2.8. Retrieved May 25, 2015, from www.gimp.org
- Traulsen, A., Pacheco, J. M., & Nowak, M. A. (2007). Pairwise comparison and selection temperature in evolutionary game dynamics. *Journal of Theoretical Biology*, *246*, 522-529.
- Traxler, M., & Gernsbacher, M. A. (2011). *Handbook of psycholinguistics*: Academic Press.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453-458.
- Tytherleigh, M. G., Bhatti, T. S., Watkins, R. M., & Wilkins, D. C. (2001). The assessment of surgical skills and a simple knot-tying exercise. *Annals of the Royal College of Surgeons of England*, *83*, 69-73.
- Ugander, J., Karrer, B., Backstrom, L., & Marlow, C. (2011). The anatomy of the facebook social graph. *arXiv preprint arXiv:1111.4503*.

- Vaesen, K. (2012). Cumulative cultural evolution and demography. *PLoS One*, 7, e40989.
- Valente, T. W. (1995). *Network models of the diffusion of innovations* (Vol. 2). Cresskill, NJ: Hampton Press.
- Vallacher, R. R., Read, S. J., & Nowak, A. (2002). The dynamical perspective in personality and social psychology. *Personality and Social Psychology Review*, 6, 264-273.
- van Schaik, C. P., Ancrenaz, M., Gwendolyn, B., Galdikas, B., Knott, C. D., Singeton, I., . . . Merrill, M. (2003). Orangutan Cultures and the Evolution of Material Culture. *Science*, 299, 102-105.
- van Schaik, C. P., & Burkart, J. M. (2011). Social learning and evolution: the cultural intelligence hypothesis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 1008-1016.
- van Schaik, C. P., Isler, K., & Burkart, J. M. (2012). Explaining brain size variation: from social to cultural brain. *Trends in cognitive sciences*, 16, 277-284.
- van Schaik, C. P., & Pradhan, G. R. (2003). A model for tool-use traditions in primates: implications for the coevolution of culture and cognition. *Journal of Human Evolution*, 44, 645-664.
- Van Vugt, M., & Ahuja, A. (2011). *Naturally Selected: Why Some People Lead, Why Others Follow, and Why It Matters*: HarperCollins.
- Wadley, L., Sievers, C., Bamford, M., Goldberg, P., Berna, F., & Miller, C. (2011). Middle Stone Age bedding construction and settlement patterns at Sibudu, South Africa. *Science*, 334, 1388-1391.
- Walker, R., Burger, O., Wagner, J., & Von Rueden, C. R. (2006). Evolution of brain size and juvenile periods in primates. *Journal of Human Evolution*, 51, 480-489.
- Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99, 5766-5771.

- Watts, Duncan J., & Dodds, Peter S. (2007). Influentials, Networks, and Public Opinion Formation. *Journal of Consumer Research*, 34, 441-458.
- Weisbuch, G., Deffuant, G., & Amblard, F. (2005). Persuasion dynamics. *Physica A: Statistical Mechanics and its Applications*, 353, 555-575.
- Wejnert, B. (2002). Integrating models of diffusion of innovations: A conceptual framework. *Annual Review of Sociology*, 28, 297-326.
- Werker, J. F., & Hensch, T. K. (2015). Critical periods in speech perception: New directions. *Annual review of psychology*, 66, 173-196.
- White, T. D., Suwa, G., & Asfaw, B. (1994). Australopithecus ramidus, a new species of early hominid from Aramis, Ethiopia. *Nature*, 371, 306-312.
- Whitehead, H., & Rendell, L. (2014). *The cultural lives of whales and dolphins*: University of Chicago Press.
- Whiten, A., & Byrne, R. W. (1988a). The Machiavellian intelligence hypothesis. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence: social complexity and the evolution fo intellect in monkeys, apes and humans* (pp. 1-9). Oxford, UK: Oxford University Press.
- Whiten, A., & Byrne, R. W. (1988b). Taking Machiavellian intelligence apart. In R. W. Byrne & A. Whiten (Eds.), *Machiavellian intelligence: social complexity and the evolution fo intellect in monkeys, apes and humans* (pp. 50-65). Oxford, UK: Oxford University Press.
- Whiten, A., & Erdal, D. (2012). The human socio-cognitive niche and its evolutionary origins. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 367, 2119-2129.
- Whiten, A., & Flynn, E. (2010). The transmission and evolution of experimental microcultures in groups of young children. *Developmental Psychology*, 46, 1694.

- Whiten, A., Hinde, R. A., Laland, K. N., & Stringer, C. B. (2011). Culture evolves. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366, 938-948.
- Whiten, A., McGuigan, N., Marshall-Pescini, S., & Hopper, L. M. (2009). *Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee* (Vol. 364).
- Whiten, A., & Van Schaik, C. P. (2007). The evolution of animal 'cultures' and social intelligence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362, 603-620.
- Wiessner, P. (2002). Hunting, healing, and hxaro exchange: A long-term perspective on! Kung (Ju/'hoansi) large-game hunting. *Evolution and Human Behavior*, 23, 407-436.
- Wood, L. A., Kendal, R. L., & Flynn, E. G. (2013). Whom do children copy? Model-based biases in social learning. *Developmental Review*, 33, 341-356.
- Wood, W., Lundgren, S., Ouellette, J. A., Busceme, S., & Blackstone, T. (1994). Minority influence: A meta-analytic review of social influence processes. *Psychological Bulletin*, 115, 323-345.
- Yong, E. (2012). Nobel laureate challenges psychologists to clean up their act. *Nature News*, 490.
- Yong, E. (2013). Psychologists strike a blow for reproducibility. *Nature*, 11, 26.
- Zou, X., Tam, K.-P., Morris, M. W., Lee, S.-l., Lau, I. Y.-M., & Chiu, C.-y. (2009). Culture as common sense: Perceived consensus versus personal beliefs as mechanisms of cultural influence. *Journal of Personality and Social Psychology*, 97, 579-597.

Appendices

Appendix A Supplementary Materials for Sociality Influences Cultural Complexity

A.1 Participants

In both experiments, participants were University of British Columbia undergraduates who completed the experiment for course credit, and monetary payments.

Experiment 1

Participants' ($N = 100$, 71 female) ages ranged between 15 and 35 ($M = 20.52$, $SD = 2.80$).

Participants were asked about their level of skill and experience with image editing software on a 7-point scale ranging from “No Experience” to “Expert”. Only participants with no experience were included in this experiment; however, due to a confusion in coding, one participant in the 1-Model treatment and one participant in the 5-Model treatment who reported 2 out of 7 in experience were included in the experiment. These participants had scores of 40/59 [1-Model] and 0/59 [5-Model], respectively, and were not outliers. Participants were randomly assigned to either the 1-Model or 5-Model treatment. We report the participant observables for each treatment in Table S1.

Table A.1. Participant observables for each treatment in Experiment 1. No differences were found between participants in each treatment.

		Treatment	
		1-Model	5-Models
Age	Mean	20.74	20.30
	SD	3.39	2.05
Gender	Males	16	13
	Females	34	37
Ethnicity	African	0	2
	East Asian	19	21
	South Asian	8	10
	Caucasian	22	15
	Mixed	1	2
Total	N	50	50

Experiment 2

Participants' ($N = 100$, 71 female) ages ranged between 17 and 37 ($M = 20.48$, $SD = 3.15$).

Participants were asked about their level of competency with rock-climbing knots and with sailing-knots on a 7-point scale ranging from “No Experience” to “Expert”. Only participants with 4 out of 7 or less on both these measures were included in this experiment and experience was controlled for. Participants were randomly assigned to either the 1-Model or 5-Model treatment. We report the participant observables for each treatment in Table S2.

Table A.2. Participant observables for each treatment in Experiment 2. No differences were found between participants in each treatment.

		Treatment	
		1-Model	5-Models
Age	Mean	20.78	20.18
	SD	2.79	3.47
Gender	Males	10	19
	Females	40	31
Ethnicity	African	2	2
	East Asian	25	25
	South Asian	6	9
	Caucasian	15	9
	Mixed	1	3
	Other	1	2
Sailing Knot Experience	Mean	1.08	1.10
	SD	0.34	0.42
Rock Climbing Knot Experience	Mean	1.14	1.16
	SD	0.40	0.55
Total	N	50	50

A.2 Experimental Design

Participants were incentivized to (a) perform to the best of their abilities and (b) transmit as much useful information as possible to the next generation. Participants were informed of these incentives with the following information: “Your performance will determine how many entries you get to win the \$100 prize. There are 5 people in each group. The top scorer in each group will get 4 entries, the second scorer will get 3 and all others will get 1. Your student is not in your group so you are not competing with them. In fact, you will receive an additional entry if your student is the best performer in their group.” Participants did not know exactly how many entries they earned, only that both their performance and that of their “students” contributed to their total entries; we actually calculated entries and ran the lottery after we had finished gathering data. In the 1-model treatment, the additional entry was given to the highest scoring participant’s model. In the 5-model treatment, we gave the additional entry to the highest scorer of the 5 potential models. Since

participants didn't know our experimental design, the existence of the other treatment, or how their transmitted information would be used, this was unlikely to have affected their incentives.

Participants showed little or no concern about the details of the incentive structure.

Both experiments consisted of learning, demonstration, and transmission. Participants first learned then demonstrated the task. During the transmission phase, participants wrote notes (Experiment 1) or created a video (Experiment 2) to assist participants in the next generation.

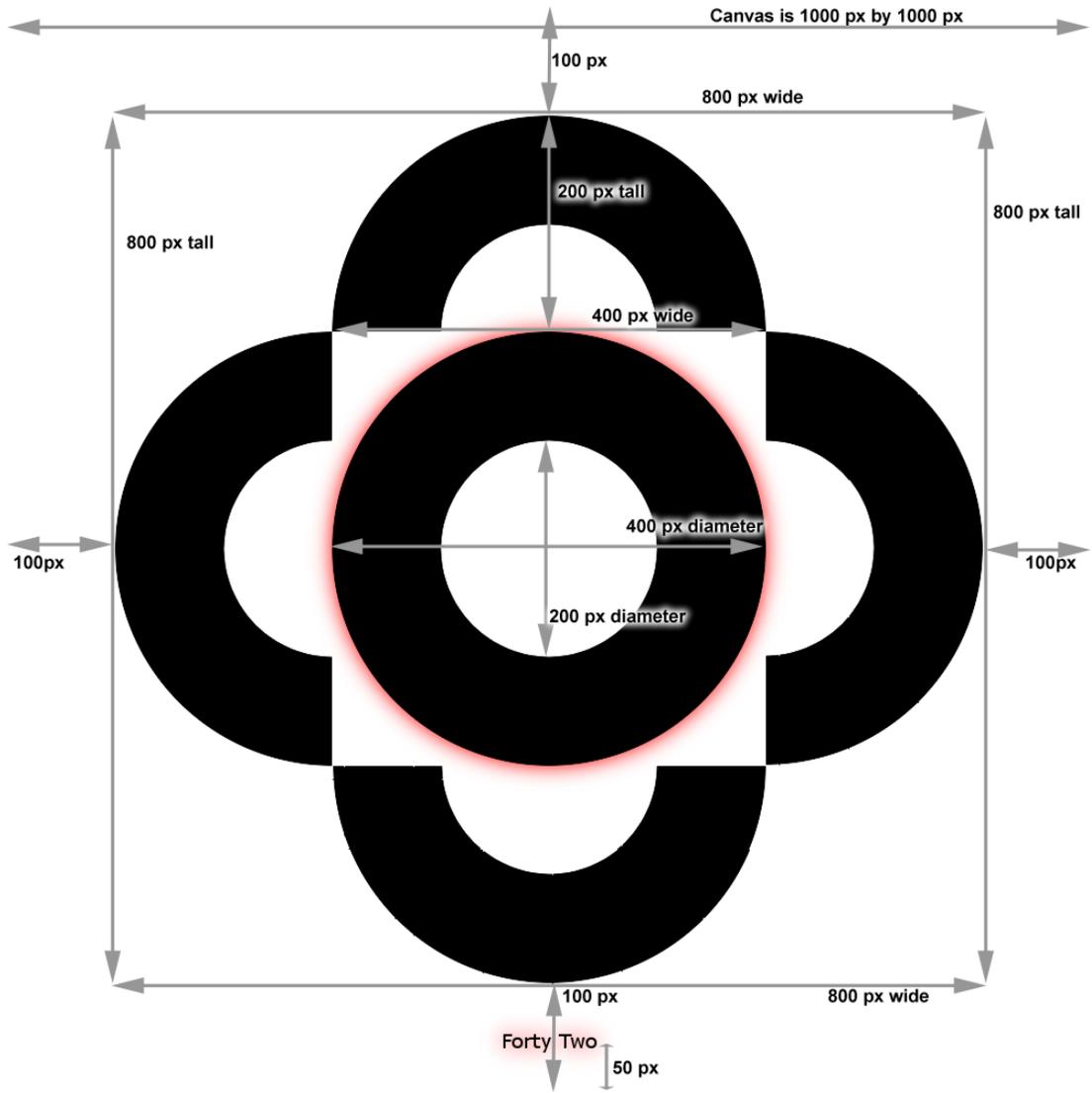


Figure A.1 Larger version of Figure 3.1b

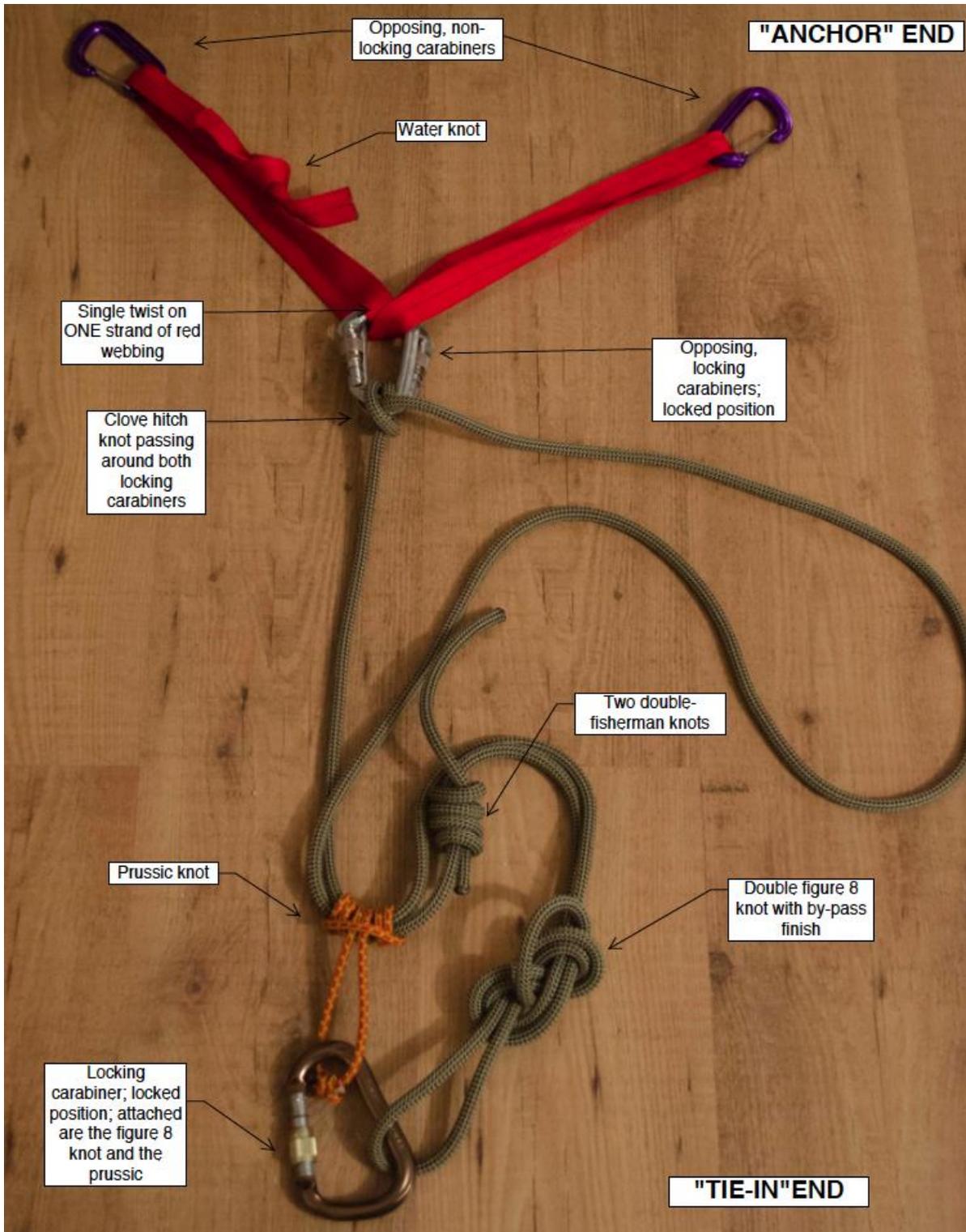


Figure A.2 Larger version of Figure 3.1c

A.3 Further Results and Analyses

Experiment 1

We estimated four OLS regression models, controlling for Age (standardized), Male (gender, with male = 1) and Ethnicity (see coding in Table S3). The model presented in the main text, with the highest adjusted R^2 , controlled for age and gender only, but the results were robust across all four models. In the first model, we used Generation, Treatment (1-Model/5-Model), and the Generation-Treatment interaction to predict the image rating. The improvement to the model by the addition of Age and Male was marginally significant, $F(94, 96) = 2.53, p = .084$. The improvement in fit by the addition of ethnicity was not large or significant and since there was a small drop in the adjusted R^2 , we chose not to report this model in the main text. We also ran a model using our controls – Age, Male, and Ethnicity – to predict the image rating. This model had a very poor fit, adjusted $R^2 = .021, p = .236$.

Table A.3. OLS regressions of standardized image rating scores. By alternating the dummy coding of treatment, we directly compare the effect of Generation by looking at the Generation coefficients.

We report several models controlling for Age, Male (gender, male = 1), and Ethnicity.

	1-Model	5-Model	1-Model	5-Model	1-Model	5-Model	1-Model/5-Model
Generation	-0.04 (0.04)	0.23 (0.04)**	-0.06 (0.04)	0.23 (0.04)**	-0.06 (0.04)	0.24 (0.04)**	
1-Model	-0.83 (0.36)*		-0.96 (0.36)**		-1.04 (0.37)**		
1-Model * Generation	0.26 (0.06)**		0.28 (0.06)**		0.29 (0.06)**		
5-Model		0.83 (0.36)*		0.96 (0.36)**		1.04 (0.37)**	
5-Model * Generation		-0.26 (0.06)**		-0.28 (0.06)**		-0.29 (0.06)**	
Age			-0.09 (0.08)	-0.09 (0.08)	-0.09 (0.08)	-0.09 (0.08)	-0.10 (0.10)
Male			-0.36 (0.19)	-0.36 (0.19)	-0.34 (0.19)	-0.34 (0.19)	-0.39 (0.23)
Ethnicity:EastAsian					-0.37 (0.62)	-0.37 (0.62)	-1.41 (0.72)
Ethnicity:SouthAsian					-0.48 (0.63)	-0.48 (0.63)	-1.26 (0.74)
Ethnicity:Caucasian					-0.57 (0.62)	-0.57 (0.62)	-1.42 (0.72)*
Ethnicity:Mixed					-1.15 (0.76)	-1.15 (0.76)	-1.72 (0.91)
(Intercept)	-0.10 (0.26)	-0.93 (0.26)**	0.11 (0.27)	-0.85 (0.26)**	0.60 (0.68)	-0.44 (0.69)	1.48 (0.71)*
R^2	0.318	0.318	0.353	0.353	0.379	0.379	0.081
Adj. R^2	0.297	0.297	0.319	0.319	0.317	0.317	0.021
N	100	100	100	100	100	100	100

** : $p < .01$, * : $p < .05$

To look for selective learning biases, we ranked the 5 potential models for participants in the 5 Models treatment from highest to lowest score (Model1 to Model5). We then broke down the sub-items in the Image Rating Scale (see section on Image Rating Scale) into 18 binary present or not present components and used the potential models' corresponding components to predict the participant's components, controlling for Generation, Age, and Male using a binary logistic regression model. This allowed us to examine how participants weighted the relative importance of their potential models. We used clustered robust standard errors (45 clusters for individuals) to control for common variance within each participant's scores. The full results (Table S4) indicate a substantial bias for learning from the best model, but also a bias for learning from the next 3 best models. To test if Models 2, 3, and 4 are indistinguishable, we ran a second binary logistic regression collapsing these models into one predictor. The fit of this model was slightly worse than the reported model ($AICc = 0.987$ vs. $AICc = 0.971$).

Table A.4. Binary logistic regression of the presence or absence of each component of the target image in each participant's attempted image on the corresponding component in each of the 5 potential models. We control for non-independence between participant's image components using clustered robust standard errors. The odds ratios reported reveal a large and significant bias for the

best model, but also biases for the 3 next best models. We control for Generation, Age (standardized), and Male (gender, male = 1).

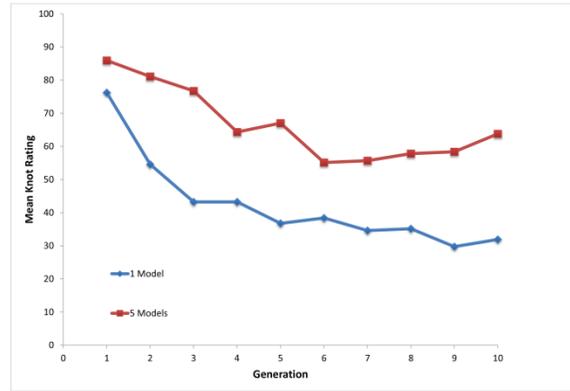
Experiment 1	
Variable	Coefficient (standard error)
Model1	3.91*** (1.26)
Model2	2.48*** (0.87)
Model3	1.75* (0.56)
Model4	2.19*** (0.58)
Model5	0.89 (0.26)
Generation	1.10 (0.09)
Age	0.95 (0.09)
Male	0.24*** (0.12)
(Intercept)	0.47 (1.01)
Pseudo R^2	0.28
N	810 (45 clusters)

Robust standard errors in parentheses

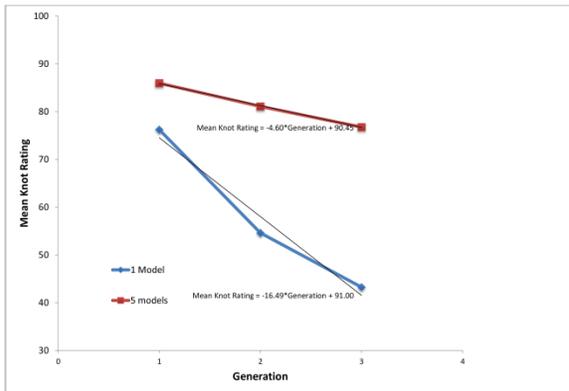
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Experiment 2

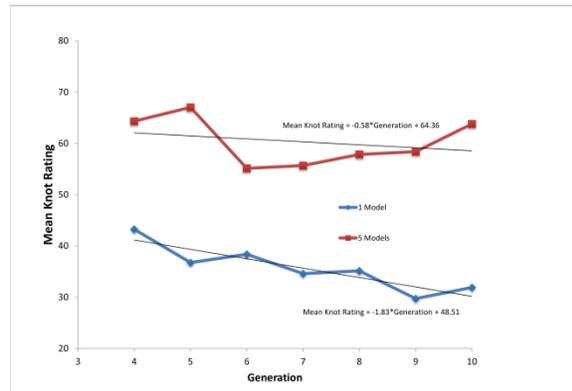
Figure S5 breaks down the differential trend in Experiment 2 between the first 3 generations (FigureS5 b) and the last 7 generations (FigureS5 c) because the first 3 generations have a more sharply declining trend. Figures S1b and S1c show that the 1-Model treatment has a much more rapid decline in skill level compared to the 5-Model treatment. In the last 7 generations, the decline is slower than in the first 3 generations in both treatments, but still more rapid in the 1 Model treatment compared to the 5 Models treatment.



(a)



(b)



(c)

Figure A.3. (a) Mean knot ratings for each generation for the 1-Model and 5-Model treatments in Experiment 2. A differential trend is visible for the earlier and later generations. (b) A closer view of the knot ratings for generations 1 to 3 for the 1-Model and 5-Model treatments. The 1-Model treatment has a more rapid decline. (c) A closer view of the knot ratings for generations 4 to 10 for the 1-Model and 5-Model treatments. The fit of the linear decline is much better for the 1-Model treatment compared to the 5-Model treatment, which shows accumulation for generations 6 to 10.

We estimated a series of OLS regression models (Table S6), controlling for Age (standardized), Male, Ethnicity, and prior experience in sailing knots and climbing knots. For each set of predictors (shown as separate tables), we estimated the model for all generations, the first 3 generations, the last 7 generations and the last 4 generations. In the first table, we used Generation, Treatment (1-

Model/5-Model), and the Generation-Treatment interaction to predict the knot rating. In the second table, we included all predictors and controls – Generation, Treatment, Generation-Treatment interaction, Age, Male, Ethnicity, and experience. In the third table, we ran a model using only our controls – Age, Male, Ethnicity, and experience. In the fourth and final table, we included all predictors and controls, except ethnicity. With the exception of the models with only the controls (third table), the models had a very similar fit. However, the model with Age, Male, and experience as controls had the highest adjusted R^2 values for the first 3 generations and was only slightly lower for the last 7 generations, so we reported the results from the models with fewer predictors (controlling for Age, Male, and prior experience) in the main text. The overall findings were robust across all models.

Table A.5. OLS regression Models of standardized knot rating scores. By alternating the dummy coding of treatment, we directly compare the effect of Generation by looking at the Generation

coefficients. We report several models controlling for Age (standardized), Male (gender, male = 1),

Ethnicity, and experience.

	1 Model:All	5 Models:All	1 Model:1-3	5 Models:1-3	1 Model:4-10	5 Models:4-10	1 Model:6-10	5 Models:6-10
Generation	-0.17 (0.04)**	-0.13 (0.04)**	-0.75 (0.26)**	-0.21 (0.26)	-0.08 (0.05)	-0.03 (0.05)	-0.08 (0.07)	0.09 (0.07)
5-Model	0.87 (0.31)**		-0.02 (0.81)		0.72 (0.52)		-0.27 (0.79)	
5-Model * Generation	0.04 (0.05)		0.54 (0.37)		0.06 (0.07)		0.17 (0.10)	
1-Model		-0.87 (0.31)**		0.02 (0.81)		-0.72 (0.52)		0.27 (0.79)
1-Model * Generation		-0.04 (0.05)		-0.54 (0.37)		-0.06 (0.07)		-0.17 (0.10)
(Intercept)	0.41 (0.22)	1.29 (0.22)**	1.65 (0.57)**	1.63 (0.57)**	-0.27 (0.37)	0.45 (0.37)	-0.28 (0.56)	-0.56 (0.56)
R^2	0.505	0.505	0.441	0.441	0.488	0.488	0.591	0.591
Adj. R^2	0.490	0.490	0.376	0.376	0.466	0.466	0.564	0.564
N	100	100	30	30	70	70	50	50

** : $p < .01$, * : $p < .05$

	1 Model:All	5 Models:All	1 Model:1-3	5 Models:1-3	1 Model:4-10	5 Models:4-10	1 Model:6-10	5 Models:6-10
Generation	-0.16 (0.04)**	-0.12 (0.04)**	-0.87 (0.37)*	-0.27 (0.30)	-0.08 (0.05)	-0.03 (0.05)	-0.04 (0.07)	0.13 (0.07)
Age	0.10 (0.08)	0.10 (0.08)	0.03 (0.18)	0.03 (0.18)	0.00 (0.10)	0.00 (0.10)	-0.01 (0.09)	-0.01 (0.09)
Male	0.36 (0.17)*	0.36 (0.17)*	0.48 (0.40)	0.48 (0.40)	0.23 (0.18)	0.23 (0.18)	0.22 (0.17)	0.22 (0.17)
Ethnicity:African	-1.27 (0.55)*	-1.27 (0.55)*	-0.46 (1.29)	-0.46 (1.29)				
Ethnicity:EastAsian	-0.69 (0.43)	-0.69 (0.43)	-1.14 (0.69)	-1.14 (0.69)	0.63 (0.37)	0.63 (0.37)	0.92 (0.36)*	0.92 (0.36)*
Ethnicity:SouthAsian	-0.85 (0.46)	-0.85 (0.46)	-0.79 (0.86)	-0.79 (0.86)	0.31 (0.39)	0.31 (0.39)	0.53 (0.40)	0.53 (0.40)
Ethnicity:Caucasian	-0.95 (0.45)*	-0.95 (0.45)*	-0.98 (0.67)	-0.98 (0.67)	0.47 (0.39)	0.47 (0.39)	0.79 (0.41)	0.79 (0.41)
Ethnicity:Mixed	-0.41 (0.58)	-0.41 (0.58)	-0.67 (1.07)	-0.67 (1.07)	0.92 (0.56)	0.92 (0.56)	0.94 (0.51)	0.94 (0.51)
Sailing knot experience	0.07 (0.08)	0.07 (0.08)	-0.04 (0.27)	-0.04 (0.27)	0.03 (0.08)	0.03 (0.08)	-0.01 (0.06)	-0.01 (0.06)
Climbing knot experience	-0.01 (0.09)	-0.01 (0.09)	0.38 (0.40)	0.38 (0.40)	-0.00 (0.08)	-0.00 (0.08)	0.01 (0.07)	0.01 (0.07)
5-Model	0.77 (0.31)*		-0.08 (0.92)		0.72 (0.52)		-0.23 (0.79)	
5-Model * Generation	0.04 (0.05)		0.60 (0.44)		0.05 (0.07)		0.17 (0.10)	
1-Model		-0.77 (0.31)*		0.08 (0.92)		-0.72 (0.52)		0.23 (0.79)
1-Model * Generation		-0.04 (0.05)		-0.60 (0.44)		-0.05 (0.07)		-0.17 (0.10)
(Intercept)	1.05 (0.45)*	1.82 (0.42)**	2.71 (1.10)*	2.63 (0.95)**	-0.87 (0.48)	-0.16 (0.52)	-1.51 (0.66)*	-1.73 (0.64)**
R^2	0.582	0.582	0.601	0.601	0.564	0.564	0.697	0.697
Adj. R^2	0.524	0.524	0.319	0.319	0.481	0.481	0.609	0.609
N	100	100	30	30	70	70	50	50

** : $p < .01$, * : $p < .05$

	1 Model:All	5 Models:All	1 Model:1-3	5 Models:1-3	1 Model:4-10	5 Models:4-10	1 Model:6-10	5 Models:6-10
(Intercept)	1.27 (0.54)*	1.27 (0.54)*	1.05 (0.66)	1.05 (0.66)	-0.69 (0.46)	-0.69 (0.46)	-0.79 (0.51)	-0.79 (0.51)
Age	0.05 (0.11)	0.05 (0.11)	0.22 (0.20)	0.22 (0.20)	-0.11 (0.12)	-0.11 (0.12)	-0.07 (0.14)	-0.07 (0.14)
Male	0.78 (0.22)**	0.78 (0.22)**	0.74 (0.47)	0.74 (0.47)	0.73 (0.22)**	0.73 (0.22)**	0.72 (0.23)**	0.72 (0.23)**
Ethnicity:African	-1.98 (0.72)**	-1.98 (0.72)**	-2.38 (1.39)	-2.38 (1.39)				
Ethnicity:EastAsian	-1.48 (0.57)**	-1.48 (0.57)**	-0.47 (0.74)	-0.47 (0.74)	0.20 (0.48)	0.20 (0.48)	0.24 (0.53)	0.24 (0.53)
Ethnicity:SouthAsian	-1.50 (0.59)*	-1.50 (0.59)*	-0.35 (0.93)	-0.35 (0.93)	0.16 (0.51)	0.16 (0.51)	0.26 (0.59)	0.26 (0.59)
Ethnicity:Caucasian	-1.77 (0.58)**	-1.77 (0.58)**	-1.41 (0.77)	-1.41 (0.77)	0.12 (0.51)	0.12 (0.51)	0.07 (0.60)	0.07 (0.60)
Ethnicity:Mixed	-0.58 (0.77)	-0.58 (0.77)	0.16 (1.24)	0.16 (1.24)	0.70 (0.74)	0.70 (0.74)	0.36 (0.77)	0.36 (0.77)
Sailing knot experience	-0.00 (0.10)	-0.00 (0.10)	0.27 (0.28)	0.27 (0.28)	-0.02 (0.10)	-0.02 (0.10)	-0.02 (0.10)	-0.02 (0.10)
Climbing knot experience	-0.07 (0.11)	-0.07 (0.11)	-0.23 (0.42)	-0.23 (0.42)	0.08 (0.11)	0.08 (0.11)	0.06 (0.11)	0.06 (0.11)
R^2	0.221	0.221	0.329	0.329	0.196	0.196	0.230	0.230
Adj. R^2	0.144	0.144	0.027	0.027	0.091	0.091	0.079	0.079
N	100	100	30	30	70	70	50	50

** : $p < .01$, * : $p < .05$

	1 Model:All	5 Models:All	1 Model:1-3	5 Models:1-3	1 Model:4-10	5 Models:4-10	1 Model:6-10	5 Models:6-10
Generation	-0.16 (0.04)**	-0.15 (0.04)**	-0.67 (0.28)*	-0.25 (0.28)	-0.07 (0.05)	-0.03 (0.05)	-0.05 (0.07)	0.11 (0.07)
Age	0.05 (0.07)	0.05 (0.07)	0.07 (0.14)	0.07 (0.14)	-0.02 (0.09)	-0.02 (0.09)	-0.00 (0.08)	-0.00 (0.08)
Male	0.35 (0.17)*	0.35 (0.17)*	0.41 (0.37)	0.41 (0.37)	0.34 (0.18)	0.34 (0.18)	0.35 (0.17)*	0.35 (0.17)*
Sailing knot experience	0.09 (0.08)	0.09 (0.08)	0.11 (0.21)	0.11 (0.21)	0.01 (0.08)	0.01 (0.08)	-0.02 (0.07)	-0.02 (0.07)
Climbing knot experience	0.04 (0.08)	0.04 (0.08)	0.19 (0.33)	0.19 (0.33)	0.07 (0.07)	0.07 (0.07)	0.04 (0.07)	0.04 (0.07)
5-Model	0.96 (0.31)**		0.28 (0.85)		0.78 (0.52)		-0.33 (0.83)	
5-Model * Generation	0.01 (0.05)		0.42 (0.38)		0.03 (0.07)		0.16 (0.10)	
1-Model		-0.96 (0.31)**		-0.28 (0.85)		-0.78 (0.52)		0.33 (0.83)
1-Model * Generation		-0.01 (0.05)		-0.42 (0.38)		-0.03 (0.07)		-0.16 (0.10)
(Intercept)	0.27 (0.23)	1.23 (0.23)**	1.40 (0.63)*	1.68 (0.65)**	-0.44 (0.38)	0.34 (0.38)	-0.57 (0.59)	-0.90 (0.60)
R^2	0.544	0.544	0.518	0.518	0.524	0.524	0.628	0.628
Adj. R^2	0.509	0.509	0.365	0.365	0.471	0.471	0.566	0.566
N	100	100	30	30	70	70	50	50

** : $p < .01$, * : $p < .05$

To test for joint significance of treatment and the treatment-generation interaction, we estimated OLS regression models with and without these predictors (Table S6). The regression model for the first 3 generations was significantly more predictive with the treatment and interaction ($R^2 = .52$) than without ($R^2 = .209$), $F(22, 24) = 7.1, p = .004$. Similarly, the regression model for the last 7 generations was significantly more predictive with the treatment and interaction ($R^2 = .52$) than without ($R^2 = .19$), $F(22, 24) = 21.6, p < .001$.

Table A.6. OLS regression models with and without the Treatment (1-Model/5-Model) and Treatment-Generation interaction used to test for joint significance. The models were significantly more predictive with the addition of these variables.

	Gen 1-3	Gen 1-3	Gen 4-10	Gen 4-10
Generation	-0.43 (0.24)	-0.67 (0.28)*	-0.04 (0.05)	-0.07 (0.05)
Age	0.05 (0.17)	0.07 (0.14)	-0.13 (0.11)	-0.02 (0.09)
Male	0.47 (0.45)*	0.41 (0.37)*	0.72 (0.21)**	0.34 (0.18)
Sailing knot experience	0.11 (0.21)	0.09 (0.08)	-0.01 (0.10)	0.01 (0.08)
Climbing knot experience	-0.06 (0.40)	0.19 (0.33)	0.14 (0.09)	0.07 (0.07)
5-Model		0.28 (0.85)		0.78 (0.52)
5-Model * Generation		0.42 (0.38)		0.03 (0.07)
(Intercept)	1.38 (0.58)*	1.40 (0.63)*	1.40 (0.63)*	-0.44 (0.38)
R^2	0.209	0.518	0.192	0.524
Adj. R^2	0.044	0.365	0.129	0.471
N	30	30	70	70

** $: p < .01$, * $: p < .05$

A.4 Normalized Cross-correlation Metric

The normalized cross-correlation metric is a pixel-based image comparison method that calculates the cross-correlation between two images, normalizing by subtracting the mean and dividing by the standard deviation:

$$NCC = \frac{1}{n} \sum_{x,y} \frac{(m(x,y) - \bar{m})(t(x,y) - \bar{t})}{\sigma_m \sigma_t}$$

Where m is the model image, t is the target image, n is the number of pixels in t and m . \bar{m} and \bar{t} are the means of m and t respectively and σ_m and σ_t are the standard deviations. The

ImageMagick similar script can be downloaded here:

<http://www.fmwconcepts.com/imagemagick/downloadcounter.php?scriptname=similar&dirname=similar>

Since a pixel-based algorithm doesn't fully account for all features in the image, we use a rating scale for analyses and the algorithm as a robustness check.

A.5 Rater Training and Testing

In Experiment 1, we trained and tested 3 raters using images from a pilot study and images from participants who had more experience than our experience threshold. We ran 4 rounds of training and testing with 7-10 images in each round, with separate images for training and testing. Inter-rater reliability was very high, $ICC(3, 1) = 0.997$.

In Experiment 2, ratings were a bit more difficult, so we trained 2 raters with knots from a pilot study and experimenter created knots that showed differences and variations in knot-tying and placement. Raters were then tested on separate knots from a pilot study. Inter-rater reliability was very high, $r = 0.87, p < 0.001$.

A.6 Rating Scales

Image Rating Scale

	Marks Possible	Marks Given	Notes – <i>part marks in italics</i>
1. Middle Black Circle			
Black circle	2		2 - for a black circle 1 - <i>for just the outline of the circle</i>
Centered	2		2 - between dotted lines
Size / Shape	2		2 - for perfect, 1 - <i>for just acceptable</i>
Hard edge – not “fuzzy” or appearing hand drawn	1		Assess ‘outer edge’ only, ‘inner edge’ is considered the edge of the middle white circle
2. Middle White Circle			
White circle	1		1 - for a white circle

	Marks Possible	Marks Given	Notes – <i>part marks in italics</i>
Centered	2		2 - between dotted lines
Size / Shape	2		2 - perfect, <i>1 - just acceptable</i>
Hard edge – not “fuzzy” or appearing hand drawn	1		
3. Half Circle “half-donuts”			
Half circle “half-donuts”	8		For each of the 4: 2 - for black half circle “half-donuts” Otherwise: <i>1 - for just the outline of the “half-donuts”</i> <i>1 - for just circles rather than “half-donuts”</i> <i>1 - for just black half circle rather than “half-donuts”</i>
Aligned	8		For each of the 4: 2 - between dotted lines
Outer Arc Size / Shape	8		For each of the 4: 2 - perfect, <i>1 - just acceptable</i>
Inner Arc Size / Shape	8		For each of the 4: 2 - perfect, <i>1 - just acceptable</i>
Hard edge – not “fuzzy” or appearing hand drawn	4		1 - For each of the 4:
4. Red Glow			
Circle	3		<i>1 - for glow appearing on circle</i> <i>1 - for glow having a gradient (fuzzy edge)</i> <i>1 - for glow having ‘regular’ shape, i.e. no ‘squiggle marks’</i>
Text	3		<i>1 - for glow appearing on text</i> <i>1 - for glow having a gradient (fuzzy edge)</i> <i>1 - for glow having ‘regular’ shape, i.e. no ‘squiggle marks’</i>
5. Text			
Forty Two	1		
Aligned	2		2 - perfect, <i>1 - just acceptable</i>
6. Extraneous Elements			
No extraneous elements	1		Extraneous elements = elements not present in the model image

Knot Rating Scale

Part/Knot	Marks Possible	Marks Given	Notes
1. Anchor - Part 1			
A. Use of correct rope and carabiners	2		2 - Correct rope and carabiners 1 - Correct carabiners and wrong rope or vice versa 0 - Both are incorrect
B. Carabiners	2		2 - Gates are opposing and pointing out 1 - Gates are pointing in 0 - Gates are not opposing
C. Water Knot	2		2 - Flat knot and correctly overlapping rope; loop is in one plane 1 - Flat knot but twist in the loop 0 - Failure to complete
2. Anchor - Part 2			
A. Twists upper strand of rope	2		2 - One twist on correct strand 1 - No twist 0 - Both strands are twisted OR neither is twisted
B. Carabiners	1		1 - Locking carabiners are chosen 0 - Incorrect carabiners (non-locking)
C. Gate Positioning and State	3		3 - Gates are opposing and locked 2 - Gates are opposing and open 1 - Gates are on the same side and locked 0 - Gates are on the same side and open
3. Clove Hitch			
A. Knot quality	3		3 - Loops overlap correctly and knot cinches tight 2 - Loops overlap incorrectly and there is no play in the knot 1 - Loops overlap incorrectly and there IS play in the knot 0 - Failure to complete
B. Position	1		1 - Knot goes around both carabiners 0 - Knot goes around one carabiner/neither
4. Figure 8			
A. Attachment	2		2 - Figure 8 knot is attached to carabiner 0 - Knot is attached to anything else (chair, loop on the chair, etc.)

Part/Knot	Marks Possible	Marks Given	Notes
B. Double Figure 8	3		3 - No cross overs in the rope Subtract mark for each cross over in the rope 0 - Knot is not a figure 8
C. Bypass finish	3		3 - Free strand is fed through the "top" of the figure 8 from the front 2 - Free strand is fed through the "top" of the figure 9 from the back 1 - Free strand is fed through the "bottom" of the figure 8 (From either the back or the front) 0 - Bypass is forgotten
5. Fisherman Knot 1			
A. Knot quality	2		2 - Two visible loops overlapping correctly 1 - 1 visible loop 0 - Failure to complete or more than 2 loops
6. Fisherman Knot 2			
A. Knot quality	2		2 - Two visible loops overlapping correctly 1 - 1 visible loop 0 - Failure to complete or more than 2 loops
7. Both Knots			
A. Rope Direction	1		1 - The two strands go in opposite directions 0 - The two strands go in the same direction
B. Clean and Tight Together?	1		1 - Fisherman's knots are pulled close together or can be slid towards each other 0 - Fisherman's knots are loosely spaced and can't be brought together
8. Prussic			
A. Creates enough wrap arounds	3		3 - 6 loops around the two ropes 2 - 4 or 8 loops around the two ropes 1 - 2 or 10 loops around the two ropes 0 - Failure to complete
B. Position	1		1 - Prussic goes around both strands of grey rope 0 - Prussic goes around only ONE strand or MORE than 2

Part/Knot	Marks Possible	Marks Given	Notes
C. Secured?	3		3 - Secured to carabiner and carabiner is locked 2 - Secured to carabiner and carabiner is open 1 - Not secured to carabiner and carabiner is locked 0 - Not secured to carabiner and carabiner is open

Appendix B Supplementary Materials for When and Who of Social Learning

B.1 Participants

All participants were recruited from the University of British Columbia's Economics Participant Pool, which is open to the public, but primarily consists of undergraduate students. In our measures and background surveys, we included 2 vigilance check questions ("Click Disagree a Little" and "Click Somewhat Agree"). Of our 101 participants, 27 failed at least one of these two checks. Table S1 reports demographics for the (a) 74 usable participants and (b) 27 who failed the vigilance check question. We were unable to predict failure using our demographics, suggesting no observable difference between the two groups. In the next section, we report our main results with the 27 excluded participants from the main text included in the analysis, revealing no substantive difference in results.

Table B.1 (a). Participant demographics for 74 participants included in main text.

		Euro	East Asian	Other	TOTAL
		Canadian	Canadian		
Age	Mean	23.95	20.69	21.46	21.73
	SD	9.57	2.59	2.77	5.55
Gender	Female	10	24	5	39
	Male	10	15	10	35

Table B.1 (b). Participant demographics for 27 participants excluded from main text.

		White Canadian	East Asian Canadian	Other	TOTAL
Age	Mean	21.75	22.60	21.00	22.30
	SD	4.19	5.38	1.00	4.86
Gender	Female	1	13	3	17
	Male	3	7	0	10

Formally defined frequency-dependent social learning strategies for N traits

To operationalize our analysis, we consider four types of formally defined frequency-dependent social learning strategies, where t is one of N cultural traits in the population, p_t is the frequency of t in the population and p_i is the probability of an individual copying t .

- (1) Conformist transmission – the disproportionate likelihood of adopting a common variant ($p_i > p_t$ if $p_t > 1/N$).
- (2) Unbiased social learning – adopting a common variant at or below the frequency of the trait in the population, but above chance ($p_t \geq p_i > 1/N$ if $p_t > 1/N$).
- (3) Asocial learning – adopting a trait independent of the population frequency ($p_i \perp p_t$, so on average $p_i = 1/N$ *ceteris paribus*).
- (4) Anti-conformity – adopting the rare trait in the population ($p_i > 1/N$ if $p_t < 1/N$).

B.2 Experimental Design

The basic experimental design is illustrated in Fig. 1 and described in the Methods section of the main text. We used an Asch-style line judgement task, which has a long history of use in psychology. Apart from comparison to past research, the task is also simple to explain, has a

uncontroversial “correct answer”, and removes priors from outside the experimental setting affecting specific decisions in the game (i.e. people don’t enter the experiment with a bias toward any particular line). Here we provide some additional details. Background measures can be found in the Background Measures section.

Experiment 1: Number of options

In Experiment 1, participants had to compare between 2 and 6 lines to identify the longest line. Each trial was worth a maximum of \$1, however, the payoff associated with each line was proportional to the length of the selected line relative to all other lines. The longest line received \$1 and the shortest line received no money. We calculated the payment (P) for each line using the following formula:

$$v = \frac{\text{max-length}}{\text{max} - \text{min}}$$
$$P = \frac{\sqrt{2 - v} - \sqrt{v}}{\sqrt{2 - v} + \sqrt{v}}$$

Where v is the relative difficulty of the line, calculated by subtracting the line length (length) from the longest line length (max) and dividing this by the difference between the longest (max) and shortest lines (min). Since v can range from 0 to 1, we can plot P over the range of v to show how the function behaves (Fig. S1 below).

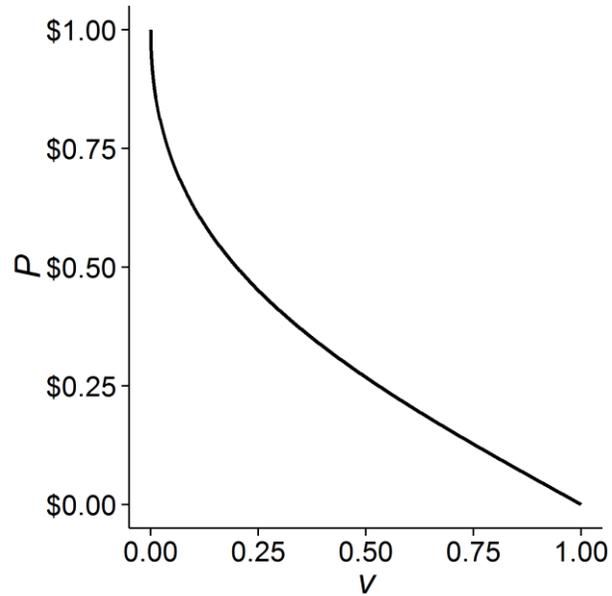


Figure B.1. Payments based on line difficulty. The function rapidly declines in payment (**P**) from the longest line ($v = 0$; $P = \$1$) and then behaves almost linearly, reducing in value to the shortest line ($v = 1$; $P = \$0$).

Experiment 2: Transmission Fidelity and Payoffs

In Experiment 2, we varied errors in transmission between 0% (only true social information) to 40% (i.e. 60% social information, 40% random). Fig. S2 below is a screenshot for how participants received this information.

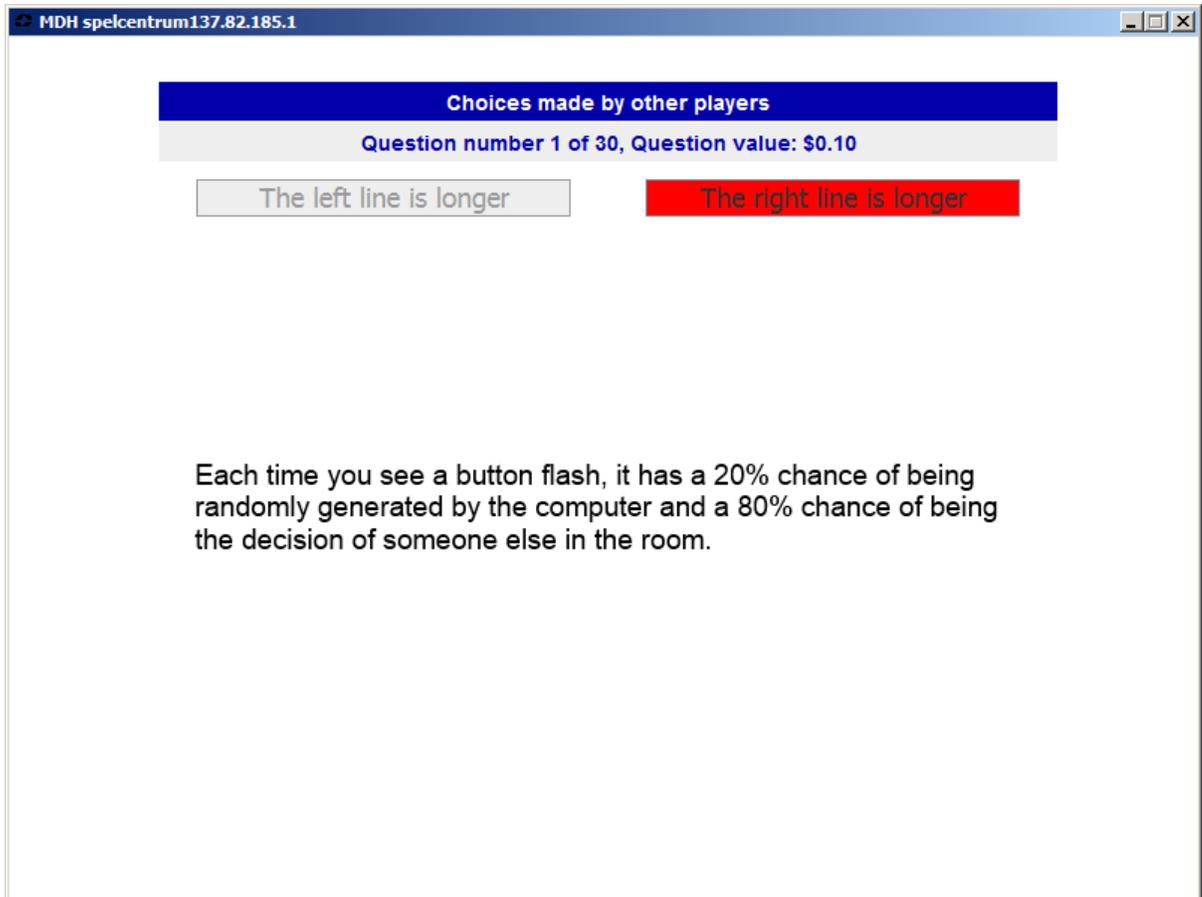


Figure B.2. Screenshot from Experiment 2 visible before and during the display of social information. Social information is conveyed in the form of flashes corresponding to the button clicked. The instructions reveal that 20% of flashes are randomly generated by the computer with the remaining 80% genuine decisions from other participants in the room.

Nakahashi, Wakano, and Henrich (2012) predictions for transmission fidelity and payoffs

The attached Mathematica file allows you to explore the effect of transmission error (migration in model) and payoffs (selection in model). We assume low selection and low error (less than 50%).

B.3 Background Measures

We measured theoretically derived individual-difference measures, other potential explanatory measures, and a variety of routine background measures. We list these below with citations, details, and sample items.

IQ: IQ was measured using Raven's Advanced Progressive Matrices (Raven & Court, 1998). Only Raven Set 1 (12 questions) was included for the first 8 participants. After this first session, we realized we had enough time to include Raven Set 2 as well, so Raven 2 was included for all other sessions. However, 3 questions were inadvertently left out of Raven 2. Although these were later added, we removed them from our analysis to maximize the number of observations. This gave us a total of 45 questions instead of 48 for combined Sets 1 and 2. There was no meaningful or significant difference in the scores for those who had these missing questions and those who did not (20.10 vs 20.08, $p = .99$). An example Raven item is shown below in Fig. S3.

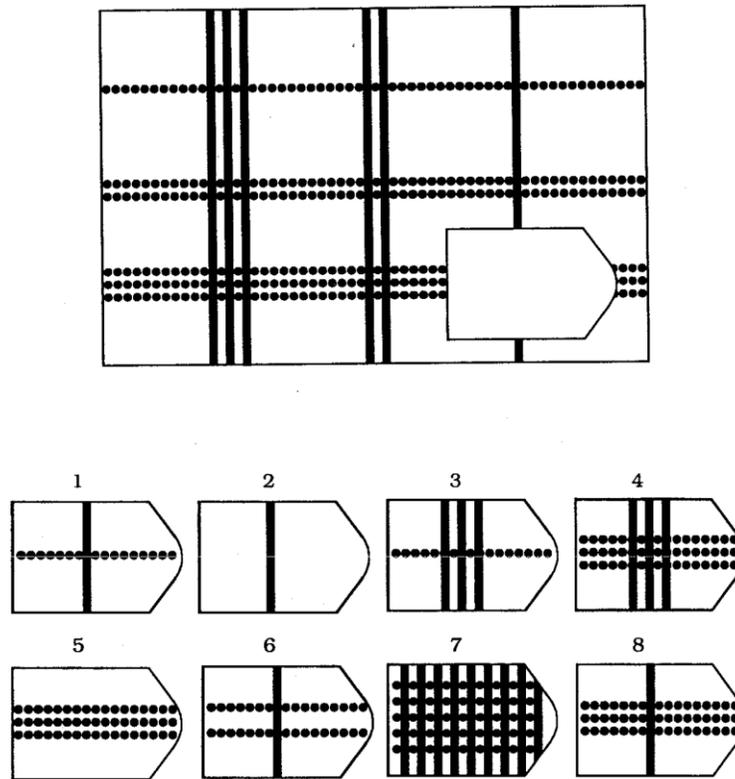


Figure B.3. Sample item from Raven's Advanced Progressive Matrices.

Prestige and Dominance: We measured self-reported prestige using the Prestige and Dominance scale (Cheng et al., 2010). An example item from the Prestige subscale: “Members of my peer group respect and admire me”. An example item from the Dominance subscale: “I enjoy having control over others”. Answers were provided using a 7-point Likert scale from “Not at all” to “Very much”.

Cultural Background: We asked for participant ethnic (or cultural) group, if they were born in Canada, how well they speak their native language, how much they identify with Canada (Inclusion of Other in the Self Scale; Aron et al., 1992), and their degree of acculturation (Vancouver Index of Acculturation; Ryder et al., 2000). We classified participant ethnicities as being East Asian Canadian, Euro-Canadian, or Other Ethnicity. The Inclusion of Other in the Self Scale involves

picking a pair of overlapping circles that best represents their level of identification with (1) their ethnic group and (2) other Canadians. The Vancouver Index of Acculturation includes Heritage Acculturation and Mainstream Acculturation subscales. An example question from the Heritage Acculturation subscale is “I often participate in my heritage cultural traditions”. An example question from the Mainstream Acculturation subscale is “I would be willing to marry a white North American person”.

Reflective vs Intuitive Thinking Styles: We measured reflective vs intuitive thinking styles using the Cognitive Reflection Test (CRT; Frederick, 2005). The CRT consists of 3 questions:

- (1) A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost? (Answer in cents)
- (2) If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets? (Answer in minutes)
- (3) In a lake, there is a patch of lily pads. Every day, the patch doubles in size. If it takes 48 days for the patch to cover the entire lake, how long would it take for the patch to cover half of the lake? (Answer in days)

Since this test is commonly used and these questions are often offered as logical puzzles, after answering the questions, we asked participants to identify any questions that they had seen before. Scores on the CRT are out of 3, so we excluded participants who had seen one or more of these questions. We also coded these questions for an Intuitive score for participants who gave the intuitive answers (some participants wrote an answer that was neither correct nor intuitive).

Personality: We measured the Big 5 Personality traits using a 44-item Big 5 Personality Inventory (John, Donahue, & Kentle, 1991; John, Naumann, & Soto, 2008). An example item from the Extraversion subscale: “I am someone who is talkative”. An example item from the

Agreeableness subscale: “I am someone who is helpful and unselfish with others”. An example item from the Conscientiousness subscale: “I am someone who does a thorough job”. An example item from the Neuroticism subscale: “I am someone who can be tense”. An example item from the Openness subscale: “I am someone who is original, comes up with new ideas”. Answers were provided using a 5-point Likert scale from “Disagree strongly” to “Agree strongly”.

Rule Following: We measured the tendency to follow rules using the Rule Following Task (RFT; Kimbrough & Vostroknutov, 2013). The RFT involves participants controlling a stick figure who walks across the screen. The stick figure stops at a series of traffic lights (screenshot shown in Fig. S4 below). Participants are told “The rule is to wait at each stop light until it turns green.” Participants were given an initial endowment of \$2.50 and each second this endowment decreased by 10c. This initial endowment was calculated such that only fully breaking the rules would ensure no loss of money. Thus participants were incentivized to break the rule and press Walk before the light turns green. How quickly they proceed in crossing the screen is a measure of their internalized rule following norms. The RFT has been shown to predict behaviour in a variety of economic games, including the public goods, dictator, ultimatum, and trust games (Kimbrough & Vostroknutov, 2013).



Figure B.4. Screenshot from the Rule Following Task.

The exact instructions provided to participants is as follows:

In the final part of this experiment, you control a stick figure that will walk across the screen.

Once the experiment begins, you can start walking by clicking the “Start” button on the left of the screen. Your stick figure will approach a series of stop lights and will stop to wait at each light. To make your stick figure walk again, click the “Walk” button in the middle of the screen.

The rule is to wait at each stop light until it turns green.

Your earnings in this part are determined by the amount of time it takes your stick figure to walk across the screen. Specifically, you begin with an initial endowment of \$2.50. Each second, this endowment will decrease by 10c and you can lose money. The game was created in Europe and says Euros, but please read these as dollars.

This is the end of the instructions for this game. If you have any questions, please raise your hand and an experimenter will answer them privately. Otherwise, please wait quietly for the experiment to begin.

If participants asked questions about the task, experimenters simply said “all instructions have been provided”.

The Rule Following Task was only included after our initial 8 participants (when we realized we had more time to include further measures) and so was included after all other measures and tasks were completed so as to ensure all participants had the same experimental experience.

Other measures: In addition to the measures discussed, we asked participants for their age, gender, degree they were enrolled in or occupation if they were not a student, major or industry, whether they had lived their entire lives in Canada, where else they had lived, what suburb they spent

most of their time in Canada, religious background, and importance of religion in their daily lives. In final debriefing questions, we also asked them to describe any strategies they were using in each game and for any remaining comments about the experiment.

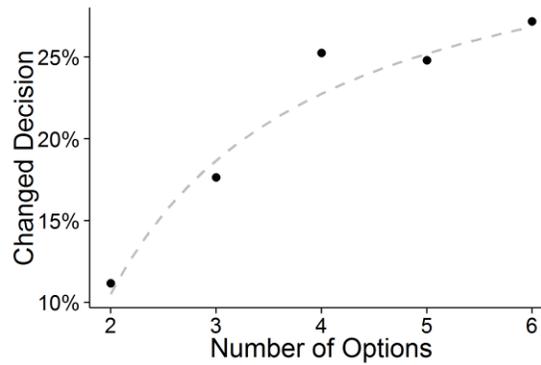
B.4 Further Analyses and Results

Here we replicate the analyses and results from the main text with the inclusion of the 27 participants who failed vigilance check questions. We also show some additional analyses mentioned in the main text.

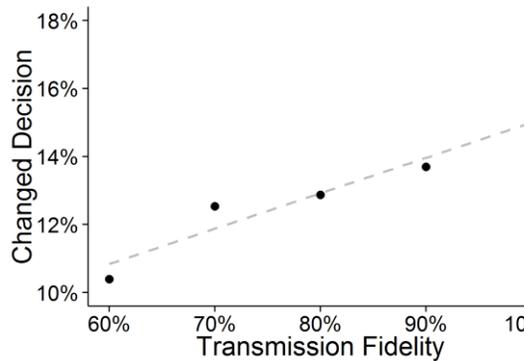
Results with all 101 participants

All tables and graphs from the main text are recreated here with the 27 exclusions included. We argue that the inclusion of these participants is defensible for the contextual predictors, since performance was incentivized. It is not defensible for individual-level predictors, which we also report, since there was no incentive to provide honest answers to these and these excluded participants failed one or more of the vigilance check questions.

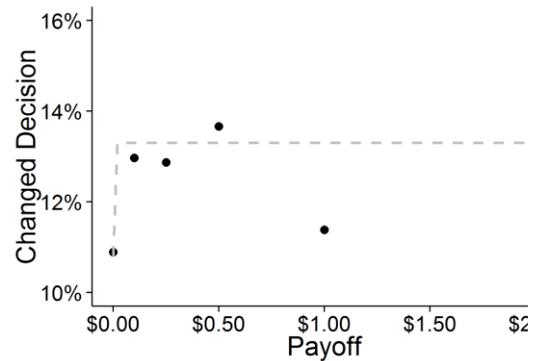
Figure 4.3 & Figure 4.5



(a)



(b)



(c)

Figure B.5. Percentage of decisions that were changed after seeing social information for (a) number of options, (b) different levels of transmission fidelity, and (c) different question payoff values.

Although there are too few points to be certain about the function that best fits these data, we used a non-linear least squares method to fit (a) to the reciprocal of traits ($y = -0.49 \cdot 1/x + 0.35$), (b) to a linear model ($y = 0.10x + 0.05$), and (c) to a step-function ($y = 0.13$ if $x > 0$; $y = 0.11$ if $x = 0$); although the pattern with \$1 and \$2 is more extreme with the inclusion of these participants. Fit functions are plotted with a grey dashed line.

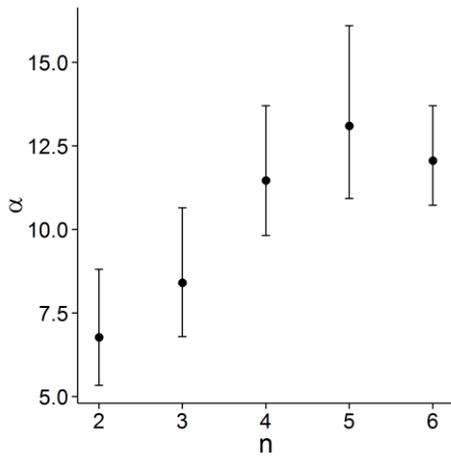
Table 4.1

Table B.2. Binary logistic multilevel model of decision to switch regressed on the proportion of participants in the option (in 10% increments for easier interpretation), the reciprocal and number of options (separate models), and the number of participants in the group. There are no substantive differences with the inclusion of the 27 participants excluded from the main analysis.

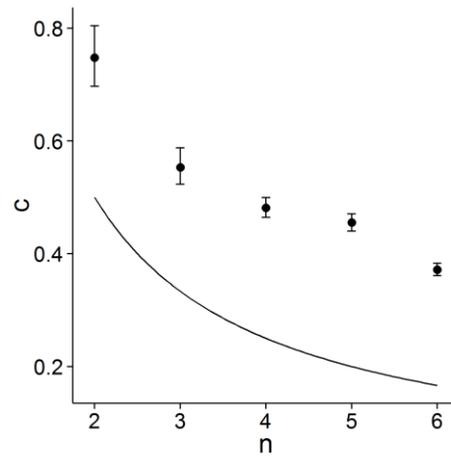
	Model 1 Number of Options	Model 2 Reciprocal of Options
(Intercept)	<0.01*** [0.00,0.01]	0.05*** [0.01,0.21]
Proportion 10%	3.05*** [2.54,3.78]	3.01*** [2.51,3.71]
Options	1.55*** [1.22,2.01]	0.08** [0.02,0.34]
Participants	1.02 [0.86,1.21]	0.99 [0.84,1.18]
AIC	261.5	263.3
Obs.	436	436
Groups	86	86

***: $p < .001$, **: $p < .01$, *: $p < .05$

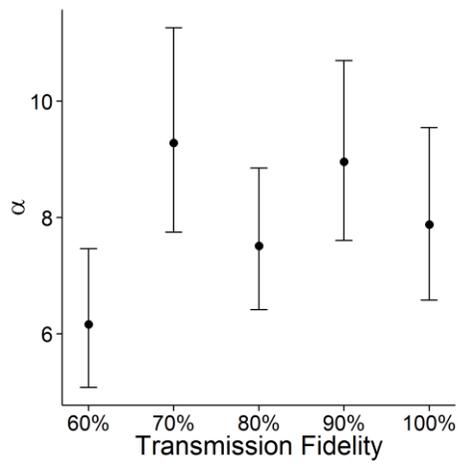
Figure 4.4 & Figure 4.6



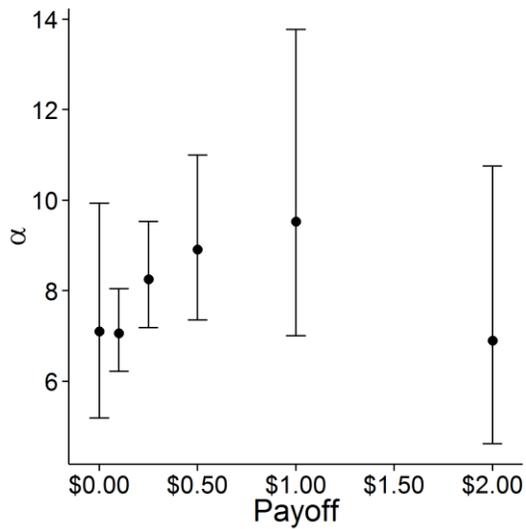
(a)



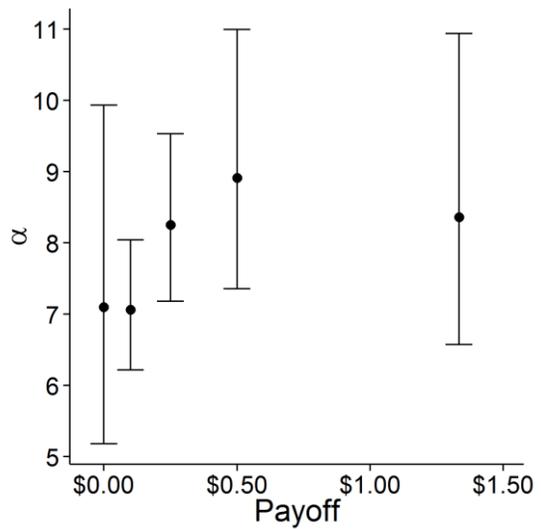
(b)



(c)



(d)



(e)

Figure B.6. Conformist bias. (a) Strength of conformist transmission parameter (α) as a function of number of options. The strength of the conformist transmission bias increases with more options. (b) Inflection point of logistic function as a function of number of options. The predicted value is shown as a solid line to distinguish it from the data (points) and model fitted values. The inflection point decreases, but remains higher than the predicted value, indicating an asocial prior. (c) Strength of conformist transmission parameter (α) as a function of transmission fidelity. Conformist transmission is strong when fidelity is higher than 60%, but at 60% it's only slightly above unbiased transmission. Strength of conformist transmission parameter (α) as a function of question payoff with (d) all payoff values and (e) \$1 and \$2 averaged to increase sample size for the highest value. The strength of the conformist transmission bias increases with diminishing returns as the payoffs increase. There are no substantive differences with the inclusion of the 27 participants excluded from the main analysis, except that there is a clearer pattern in (a) for an increased conformist bias with more traits.

Table 4.2

Table B.3. Binary logistic multilevel model of decision to switch to majority on majority size, transmission fidelity, payoff, and number of participants in the group. All coefficients are odds ratios. We control for common variance created by multiple observations from the same person with

random effects for each individual. There are no substantive differences with the inclusion of the 27 participants excluded from the main analysis.

	Model 1 Linear Payoff 10c	Model 2 Binary Payoff
(Intercept)	0.88 [0.59,1.30]	0.39* [0.16,0.93]
Majority 10%	3.18*** [2.59,3.90]	3.18*** [2.63,3.84]
Transmission 10% Payoff	1.27*** [1.11,1.46]	1.27*** [1.11,1.45]
Participants	1.02 [0.97,1.08]	2.31* [1.03,5.18]
	1.25** [1.07,1.47]	1.25** [1.06,1.47]
AIC	1017.6	1005.9
Obs.	1114	1114
Groups	101	101

***: $p < .001$, **: $p < .01$, *: $p < .05$

Table 4.3

These results should be treated with caution, since it includes individual-difference measures from those who failed one or more vigilance check questions and may therefore have entered nonsense data for other individual-difference measures.

Table B.4. OLS regression model percentage of decisions that were changed after viewing social information regressed on theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the reference group. These results show a negative relationship between IQ and social learning with higher IQ resulting in less social learning. The regression models reported show all theoretically inspired predictors; the regression model is significant when the non-significant predictors are removed (see Reduced Model below). Unsurprisingly, with the addition of those who failed the vigilance check, IQ is no longer a significant predictor in Experiment 2.

	Experiment 1	Experiment 2
(Intercept)	0.20 (0.14,0.27)***	0.21 (0.07,0.34)**
zIQ	-0.03 (-0.06,-0.00)*	-0.02 (-0.03,-0.00)
zPrestige	-0.02 (-0.05,0.02)	-0.02 (-0.04,0.01)
East Asian	0.01 (-0.06,0.08)	-0.01 (-0.06,0.03)
Other Ethnicity	-0.01 (-0.10,0.08)	0.04 (-0.02,0.09)
zAge	-0.01 (-0.04,0.02)	-0.01 (-0.02,0.01)
Male	0.02 (-0.04,0.08)	-0.01 (-0.03,0.04)
R^2	0.08	0.10
N	101	101

***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$

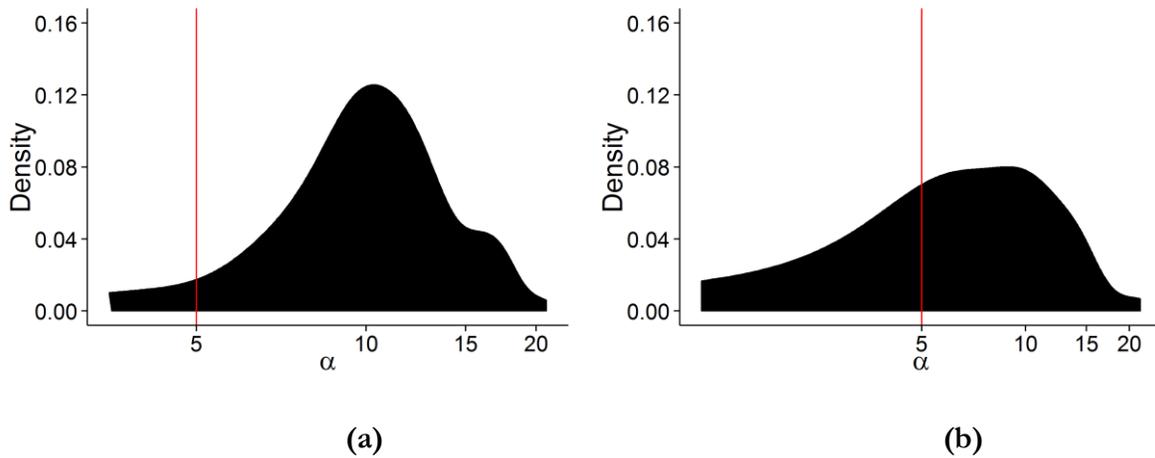


Figure B.7. Density distribution of α conformist transmission values in (a) Experiment 1 and (b) Experiment 2, with α calculated after scaling frequency of options by transmission fidelity. The red line indicates the cut off for conformist transmission with values to the left of this line indicating unbiased social learning. The x-axis is log-scaled. For visual purposes, we remove some outliers – see Density Plot with Outliers for figure including these.

Table 4

These results should be treated with caution, since it includes individual-difference measures from those who failed one or more vigilance check questions and may therefore have entered nonsense data for other individual-difference measures.

Table B.5. OLS regression model of standardized log measures of strength of conformist transmission (α) regressed on our theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the reference group. These results suggest a consistent quadratic (U shaped) relationship between IQ and the strength of the conformist transmission bias. Both those who score high or very low on the IQ test are more likely to have stronger conformist transmission biases than those who score in the middle. In Experiment 1, which is arguably more sensitive than Experiment 2 because there are often more than 2 options, conformist biases strengthen among older individuals. In Experiment 2, we were unable to fit a sigmoid to the decisions of one of the individuals who failed a vigilance check question.

	Experiment 1	Experiment 2
(Intercept)	-0.27 (-0.70,0.17)	0.21 (-1.33,1.76)
zIQ	0.02 (-0.19,0.24)	0.07 (-0.17,0.30)
zIQ ²	0.21 (0.07,0.35)**	0.18 (0.03,0.33)*
zPrestige	-0.09 (-0.29,0.11)	-0.04 (-0.32,0.23)
East Asian	0.21 (-0.26,0.68)	-0.37 (-0.88,0.14)
Other Ethnicity	0.37 (-0.20,0.94)	0.01 (-0.60,0.63)
zAge	0.30 (0.11,0.49)**	-0.13 (-0.33,0.08)
Male	-0.28 (-0.66,0.10)	0.09 (-0.31,0.50)
<i>R</i> ²	0.23***	0.11
N	101	100

***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$

Reduced Model for Social Learning with Just IQ

Table B.6. OLS regression model percentage of decisions that were changed after viewing social information regressed on IQ. These results show a negative relationship between IQ and social learning with higher IQ resulting in less social learning. The models are significant in both experiments.

	Experiment 1	Experiment 2
(Intercept)	0.22 (0.19,0.26)***	0.13 (0.11,0.15)***
zIQ	-0.04 (-0.07,-0.01)*	-0.02 (-0.03,-0.00)*
R^2	0.08*	0.05*
N	74	74

***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$

Analyses with age and gender

Here we show the full models controlling for age and gender for the contextual variables: number of options (Table 1 in main text) and transmission fidelity and payoff (Table 2 in main text).

Table 1

Table B.7. Binary logistic multilevel model of decision to switch regressed on the proportion of participants in the option (in 10% increments for easier interpretation), the reciprocal and number of options (separate models), and the number of participants in the group. All coefficients are odds ratios. We control for common variance created by multiple observations from the same person with random effects for each individual.

	Model 1 Number of Options	Model 2 Reciprocal of Options
(Intercept)	<0.01*** [0.00,0.05]	0.16 [0.01,2.51]
Proportion 10%	3.71*** [2.80,4.92]	3.64*** [2.76,4.81]
Options	1.74*** [1.25,2.42]	0.04** [0.00,0.31]
Participants	0.93 [0.74,1.18]	0.90 [0.71,1.14]
Age	0.95 [0.88,1.02]	0.95 [0.89,1.02]
Male	1.55 [0.66,3.68]	1.42 [0.61,3.33]
AIC	173.4	175.2
Obs.	332	332
Groups	64	64

***: $p < .001$, **: $p < .01$, *: $p < .05$

Table 4.2

Table B.8. Binary logistic multilevel model of decision to switch to majority on majority size, transmission fidelity, payoff, and number of participants in the group. All coefficients are odds ratios. We control for common variance created by multiple observations from the same person with random effects for each individual.

	Model 1 Linear Payoff 10c	Model 2 Binary Payoff
(Intercept)	1.45 [0.36,5.82]	0.54 [0.10,2.86]
Majority 10%	3.57*** [2.79,4.58]	3.58*** [2.78,4.61]
Transmission 10%	1.31*** [1.12,1.53]	1.31*** [1.12,1.53]
Payoff	1.03 [0.97,1.09]	2.57 ⁺ [0.99,6.69]
Participants	1.27** [1.06,1.53]	1.27* [1.05,1.53]
Age	0.99 [0.93,1.05]	0.99 [0.93,1.05]
Male	1.37 [0.67,2.78]	1.41 [0.70,2.82]
AIC	763.7	753.1
Obs.	818	818
Groups	74	74

***: $p < .001$, **: $p < .01$, *: $p < .05$

Results for Experiment 2 without Scaling

In Experiment 2, to calculate α conformist bias scores for each participant, we scaled the proportion of participants for each option by the transmission fidelity, since we knew that there was a linear relationship between transmission fidelity and social learning (Fig. 5a). Here we report the results without this scaling.

Figure 4.7

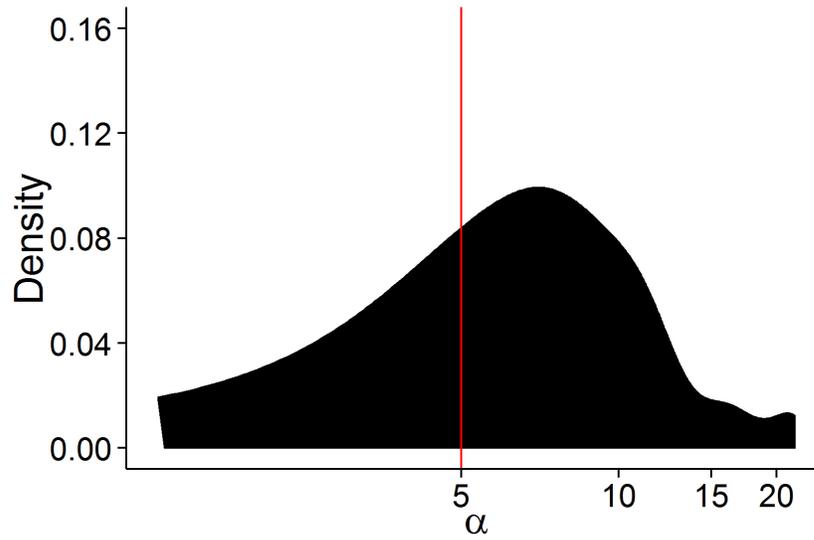


Figure. B.8. Density distribution of α conformist transmission values in Experiment 2, with α unscaled by transmission fidelity. The red line indicates the cut off for conformist transmission with values to the left of this line indicating unbiased social learning. The x-axis is log-scaled. For visual purposes, we remove some outliers.

Table 4.4

Table B.9. Experiment 2 OLS regression model of standardized log measures of strength of conformist transmission (α) regressed on our theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the reference group. The model has a worse fit without scaling by transmission fidelity, but the overall pattern remains the same.

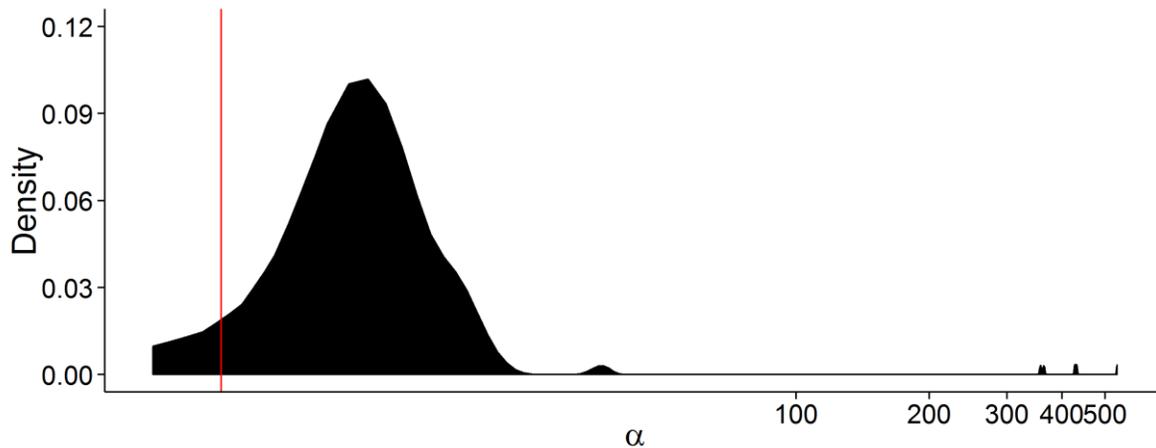
	Unscaled Experiment 2	Scaled Experiment 2
(Intercept)	<0.01 (-1.64,2.29)	0.33 (-1.64,2.29)
zIQ	0.08 (-0.18,0.37)	0.09 (-0.18,0.37)
zIQ ²	0.22 (0.09,0.42)**	0.25 (0.09,0.42)**
zPrestige	-0.03 (-0.45,0.26)	-0.10 (-0.45,0.26)
East Asian	-0.20 (-0.77,0.36)	-0.20 (-0.77,0.36)
Other Ethnicity	<0.01 (-0.60,0.75)	0.07 (-0.60,0.75)
zAge	-0.03 (-0.29,0.20)	-0.05 (-0.29,0.20)
Male	<0.01 (-0.43,0.51)*	0.04 (-0.43,0.51)
R^2	0.12	0.17 ⁺
N	74	74

***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$

Density plot with outliers

Shown below are the density plots with outliers for conformist bias scores. Even after transforming there remain some outliers. We show our results are robust to these outliers by replicating the analysis reported in Table 4 using a Robust Linear Model.

Figure 4.7 with outliers



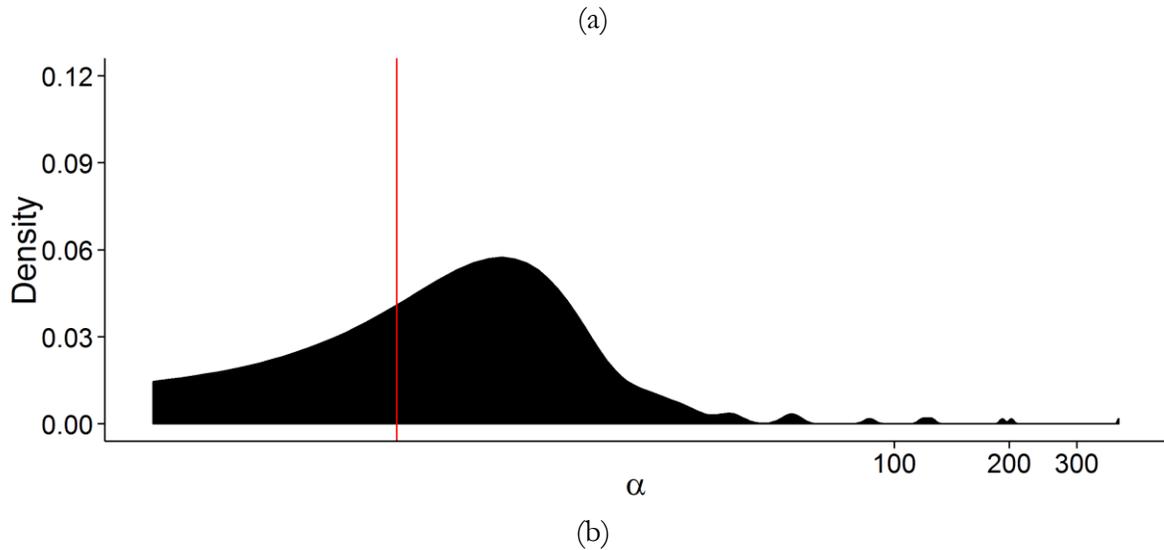


Figure B.9. Density distribution of α conformist transmission values in (a) Experiment 1 and (b) Experiment 2, with α calculated after scaling frequency of options by transmission fidelity. The red line indicates the cut off for conformist transmission with values to the left of this line indicating unbiased social learning. The x-axis is log-scaled.

Robust Linear Model Conformist Bias Analysis (Table 4)

To deal with outliers, we calculate the robust linear model using “rlm” from the “MASS” R package, which uses an MM-type regression estimator (Koller & Stahel, 2011). The general U-shaped relationship between IQ and the strength of the conformist-bias remains the same.

Table B.10. Robust linear regression model of standardized log measures of strength of conformist transmission (α) regressed on our theoretical predictors as well as age and gender. All predictors with a “z” prefix are standardized z-scores. Ethnicity was dummy coded, with Euro Canadians as the reference group. These results suggest a consistent quadratic (U shaped) relationship between IQ and the strength of the conformist transmission bias. Both those who score high or very low on the IQ test are more likely to have stronger conformist transmission biases than those who score in the middle. In Experiment 1, which is arguably more sensitive than Experiment 2 because there are often more than 2 options, conformist biases strengthen among older individuals.

	Experiment 1			Experiment 2		
	<i>b</i>	Std. Err.	<i>t</i> value	<i>b</i>	Std. Err.	<i>t</i> value
(Intercept)	-0.29	0.11	-2.67	-0.42	0.78	-0.55
zIQ	0.08	0.06	1.30	0.02	0.11	0.16
zIQ ²	0.13	0.03	3.66	0.30	0.07	4.57
zPrestige	-0.04	0.05	-0.76	0.03	0.14	0.20
East Asian	0.11	0.12	0.91	-0.11	0.27	-0.48
Other Ethnicity	0.15	0.14	1.08	0.08	0.27	0.29
zAge	0.17	0.05	3.26	-0.09	0.10	-0.98
Male	-0.22	0.10	-2.19	-0.03	0.19	-0.14
N		74			74	

Performance

Here we show that our individual predictors do not predict performance. However, people do improve after social information, but the improvement is small (approximately 3% in for both experiments).

Predicting performance

No individual predictor was particularly effective (we tried several reduced models). With all individual predictors, we were still only able to explain 16% and 13% of the variance of the asocial decision and 8% and 13% of the social decision. In the analyses reported below, we omit the Cognitive Reflection Test (both Reflective and Intuitive scores) and the Rule Following Test, since some participants had seen some of the questions in the former and some early participants did not perform the latter. However, analyses with these included suggested that these did not reliably or significantly predict performance.

Asocial Decision

Table B.11. Standardized asocial score regressed on all individual-level predictors. The model is not significant, nor are any predictors. Reduced models are also not significantly predictive.

	Experiment 1	Experiment 2
(Intercept)	0.27 (−0.26,0.81)	0.14 (−0.41,0.69)
zIQ	0.09 (−0.18,0.36)	0.20 (−0.08,0.47)
zPrestige	−0.03 (−0.34,0.28)	0.15 (−0.17,0.46)
East Asian	−0.29 (−0.91,0.34)	−0.20 (−0.84,0.43)
Other Ethnicity	−0.11 (−0.83,0.62)	0.09 (−0.65,0.82)
zAge	−0.13 (−0.40,0.14)	−0.10 (−0.37,0.18)
Male	−0.21 (−0.74,0.31)	−0.11 (−0.65,0.43)
zDominance	0.20 (−0.13,0.53)	−0.15 (−0.48,0.19)
zOpenness	−0.21 (−0.74,0.31)	−0.00 (−0.29,0.28)
zConscientiousness	0.12 (−0.15,0.39)	0.02 (−0.26,0.30)
zExtraversion	−0.31 (−0.63,0.00)	0.10 (−0.22,0.42)
zAgreeableness	0.01 (−0.29,0.32)	−0.10 (−0.41,0.21)
zNeuroticism	−0.19 (−0.52,0.14)	0.19 (−0.15,0.52)
R^2	0.16	0.13
N	74	74

***: $p < .001$, **: $p < .01$, *: $p < .05$, +: $p < .1$

Social Decision

Table B.12. Standardized social score regressed on all individual-level predictors. The model is not significant, nor are any predictors. Reduced models are also not significantly predictive.

	Experiment 1	Experiment 2
(Intercept)	0.05 (−0.51,0.61)	0.32 (−0.23,0.87)
zIQ	−0.01 (−0.29,0.27)	0.11 (−0.17,0.38)
zPrestige	−0.04 (−0.37,0.28)	0.07 (−0.25,0.38)
East Asian	−0.14 (−0.79,0.52)	−0.55 (−1.19,0.09) ⁺
Other Ethnicity	−0.12 (−0.64,0.87)	−0.26 (−0.99,0.48)
zAge	−0.09 (−0.37,0.19)	−0.15 (−0.42,0.13)
Male	−0.00 (−0.55,0.55)	0.05 (−0.49,0.58)
zDominance	−0.05 (−0.39,0.29)	0.16 (−0.49,0.18)
zOpenness	0.20 (−0.10,0.49)	−0.13 (−0.42,0.15)
zConscientiousness	0.07 (−0.22,0.35)	0.12 (−0.16,0.40)
zExtraversion	−0.18 (−0.51,0.15)	−0.02 (−0.34,0.30)
zAgreeableness	−0.15 (−0.47,0.17)	−0.12 (−0.43,0.19)
zNeuroticism	−0.13 (−0.47,0.22)	0.15 (−0.19,0.48)
<i>R</i> ²	0.08	0.13
N	74	74

Based on effect sizes, across both experiments and asocial and social decisions, IQ seems to positively predict performance (apart from Experiment 1 social decision, where it isn't predictive).

Being East Asian or older appears to negatively predict performance. Status and personality are not reliably predictive. However, all these effect sizes are statistically indistinguishable from zero.

Asocial vs Social Decision

We conducted a paired sample Student's t-test for percentage scores before and after receiving social information. In Experiment 1 with multiple options, there was a marginally significant ~3% improvement (51.2% vs 54.1%, $t(73) = -2.00, p = .050$). In Experiment 2 with only two options, there was a significant ~3% improvement (65.0% vs 68.5%, $t(73) = -4.94, p < .001$).